DESIGN AND SYNTHESIS OF SELECTIVE INHIBITORS FOR β-CATENIN/T-CELL FACTOR-4 PROTEIN-PROTEIN INTERACTION

by

Zheng Huang

A dissertation submitted to the faculty of
The University of Utah
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Department of Chemistry

The University of Utah

August 2017

The University of Utah Graduate School

STATEMENT OF DISSERTATION APPROVAL

The dissertation of **Zheng Huang**

has been approved by the following supervisory committee members:

Jon D. Rainier	, Chair	**August 8, 2016** Date Approved
Matthew S. Sigman	, Member	**August 8, 2016** Date Approved
Vahe Bandarian	, Member	**August 8, 2016** Date Approved
Cynthia J. Burrows	, Member	**August 8, 2016** Date Approved
Thomas Cheatham	, Member	**August 8, 2016** Date Approved

and by **Cynthia J. Burrows**, Chair/Dean of

the Department/College/School of **Chemistry**

and by David B. Kieda, Dean of The Graduate School.

ABSTRACT

Selective disruption of protein-protein interactions by small molecules is important for probing the structure and dynamic aspects of the cellular network. It can also provide new therapeutic targets. β-Catenin of the canonical Wnt signaling pathway uses the same positively charged groove to bind with T-cell factor (Tcf), cadherin, and adenomatous polyposis Coli (APC). The extravagant formation of β-catenin/Tcf interactions drives the initiation and progression of many cancers and fibroses, while β-catenin/cadherin and β-catenin/APC interactions are essential for cell–cell adhesion and β-catenin degradation. In this study, a selective binding site that can differentiate β-catenin/Tcf, β-catenin/cadherin and β-catenin/APC interactions was identified by alanine scanning and biochemical assays.

In Chapter 2, a new peptidomimetic strategy was used to design selective small-molecule inhibitors for β-catenin/Tcf interactions. A potent inhibitor was discovered to bind with β-catenin and completely disrupt β-catenin/Tcf interactions. It also exhibits dual selectivity for β-catenin/Tcf over β-catenin/cadherin and β-catenin/APC interactions in both biochemical and cell-based assays. The study provides a proof of concept for designing selective inhibitors of β-catenin/Tcf interactions.

To develop a more drug-like compound, Chapter 3 focused on the modification of the inhibitors discovered in Chapter 2. The replacement of the peptide bond with a benzene ring brought up an interesting off-target outcome.

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

[α]	Specific rotation
Ac	acetyl
APC	Adenomatous Polyposis Coli
BCL9	B-cell CLL/lymphoma 9
Bn	benzyl
Boc	*tert*-butyloxycarbonyl
CBP	CREB-binding protein
Cdc73	Cell Division Cycle 73
CK1α	casein kinase 1α
conc.	concentrated
CRT	β-catenin-regulated transcription
d	doublet (in NMR)
DMAP	4-dimethylaminopyridine
DMF	*N*,*N*-dimethylformamide
DMSO	dimethyl sulfone
Dvl	disheveled
EDC	1-Ethyl-3-(3-dimethylaminopropyl)carbodiimide
ELISA	enzyme-linked immunosorbent assay

ESI	electron spray ionization
Et	ethyl
Fmoc	Fluorenylmethyloxycarbonyl
FP	fluorescence polarization
Fz	fizzled
g	gram(s)
GSK3	glycogen synthase kinase 3
GSK3β	glycogen synthase kinase 3β
h	hour(s)
HEK	human embryonic kidney cell
HOAt	1-hydroxy-7-azabenzotriazole
HRMS	High-resolution mass spectrometry
HTS	high-throughput screening
Hz	hertz
IBCF	*iso*-butyl chloroformate
IC50	half maximal inhibitory concentration
*i*Pr	*iso*-propyl
ITC	Isothermal Titration Calorimetry
J	coupling constant (in NMR)
K_A	association constant
K_d	dissociation constant
K_i	enzyme-inhibitor constant
Lef	lymphoid enhancer factor

LRP	low-density lipoprotein receptor-related protein
M	moles per liter
m	multiplet (in NMR)
m/z	mass to charge ratio (in mass spectrometry)
*m*CPBA	*m*-chloroperoxybenzoic acid
MCSS	multiple copy simultaneous search
Me	methyl
MHz	megahertz
MIFs	molecular interaction fields
min	minutes
MMTV	Mouse Mammary Tumor Virus
mol	mole(s)
MS	mass spectrometry
NBS	*N*-bromosuccinimide
nM	Nanomolar
NMM	*N*-methyl morpholine
NMR	nuclear magnetic resonance
NSAIDS	non-steroidal anti-inflammatory drugs
PDB	protein databank
Ph	phenyl
PPIs	protein—protein interactions
ppm	parts per million (in NMR)
Py	Pyridine

q	quartet (in NMR)
quant.	quantitative
rt	room temperature
s	singlet (in NMR)
SAR	structure-activity relationship
SD	standard deviation
shRNAs	short hairpin RNA
siRNAs	Small interfering RNA
SPR	Surface plasmon resonance
t	triplet (in NMR)
TBAF	tetrabutylammonium fluoride
TBS	*tert*-butyldimethylsilyl
*t*Bu	*tert*-butyl
*t*BuOK	*tert*-butyl oxide
Tcf	T-cell factor
TFA	Trifluoroacetic acid
TLC	thin layer chromatography
TMS	trimethyl silyl
Tol	toluene
TR-FRET	time-resolved fluorescence resonance energy transfer
TsOH	*p*-toluenesulfonic acid
β-TRCP	*β*-transducin repeats containing proteins
μM	micromolar

ACKNOWLEDGEMENTS

First I would like to thank my committee members Dr. Burrows, Dr. Sigman, Dr. Bandarian, Dr. Cheatham and especially Dr. Rainier, for giving me tremendous support and having patience with me. I have not been a good student, always turning in my reports late and with countless mistakes, but they always see good things out of me, and when I already gave up, they gave me the courage and confidence to keep moving forward.

Looking back to the time I spent in the lab, I am thankful to have worked with Dr. Ji for the past 5 years. I did learn a lot from him and his wife, Dr. Grace Zhang. Their determination, diligence, and passion for research are something I will always look up to.

I am thankful for having worked with an amazing crew of the Ji group. I could not ask for better coworkers and friends like Levon Katsakhyan, Jack Wisniewski, and Toni Guo. Specially, I would like to thank Vanja Panić, who always patiently explained lots of biology to me and helped me with writing. I would also like to thank all the former and current group members who contributed to my projects. I would like to thank Leon Catrow for helping me with the computational part of my research, Levon Katsakhyan and Shawn Burton for the help of the first project, and Kevin Teuscher for helping with the second project. Especially, I would like to thank the two postdoc fellows Dr. Binxun Yu and Dr. Yongqiang Zhang. Dr. Yu gave me numerous suggestions and tricks in organic synthesis, and Dr. Zhang enlightened me in medicinal chemistry as well as guidance for life. I also

would like to thank everyone who has helped me in the chemistry department: Jo Vallejo, Dr. Jim Miller, Dr. Peter Flynn and Dr. Atta Arif.

Moving abroad all by myself was a tough life transition. Luckily, I have made lots of friends here, who always gave me the support and help I need, and because of them, I no longer feel alone. Firstly I would like to thank Zach Chan, Daimei Zhu and Kyan Wei for being like my family here. I would like to thank my best friend Giusseppe Huaman, who always believes in me and pushes me to do my best during my toughest time. I am glad to have Tej Singh as my friend, too, who always motivates me in volleyball as well as in life. I would also like to thank my friend for 12 years, Shuying Fan, for the joyful memories in life and also the great help and support during my thesis writing.

Finally, I thank my family for the unconditional support and love. This work is dedicated to my parents, who gave me the best they could, and always encourage me to go further, climb higher, and reach for better.

CHAPTER 1

INTRODUCTION

Each year in the United States, almost 150,000 patients are diagnosed with colorectal cancer, causing more than 55,000 deaths.[1] Current therapy for colon cancer includes rarely curative surgical resection and traditional cytotoxic agents (chemotherapy), which also have very limited effects.[2] Thus, it is important to develop new treatments taking the pathogenic molecular lesions into consideration.

1.1 History of Wnt Signaling Pathway

The Wnt signaling pathway was discovered over 30 years ago, but it can be traced to as early as the 1930s when mouse models of cancer and oncogenic retroviruses were studied.[3] In 1936, Bittner and Korteweg separately reported that some laboratory mice are more prone to acquire breast cancer, and the disease is often transmitted to their off-springs through milk.[4-5] This tumor-generating factor in the milk was not isolated until 1962 and was discovered to be a retrovirus—Mouse Mammary Tumor Virus (MMTV).[6] The studies of oncogenic retroviruses led to the discovery of viral oncogenes such as *Src*, *Myc* and *Ras*.[7] From the 1980s to the 1990s, numerous new discoveries were made to elucidate the relationship between these proto-oncogenes and cancer. A timeline for the history of *int1*, *wingless* genes and some components of the Wnt signaling pathway is shown in **Table 1.1**.

Table 1.1. Timeline of the discovery of the int1 and wnt genes.

Year	Event	Authors
1976	*Wingless* gene was identified.	Sharma and Chopra[8]
1980	Segment polarity genes were defined (*Wingless*, *Armadillo*, and *Arrow*).	Nusslein-Volhard and Wieschaus[9]
1982	*Int1* gene was identified.	Nusse and Varmus[10]
1984	Structure and sequence of the *int1* gene were disclosed.	Van Ooyen and Nusse[11]
1984	*Int2* was discovered.	Dickson *et al.*[12]
1985	Proto-oncogenes were discovered.	Bishop and Varmus[13]
1985	cDNA of the *int1* gene was elucidated-no homology with other genes at that time.	Fung *et al.* [14]
1985	Human *int1* is 99% identical to that of the mouse homolog.	Van Ooyen *et al.*[15]
1986	Overexpression of *int1* rarely leads to the formation of cells that can turn into tumor.	Brown *et al.*[16]
1987	*Int1* was found close to *wingless* and played a role in segmentation.	Rijsewijk *et al.*[17]
1987	*Wingless* gene cloned was identical to the Drosophila *int1* clone.	Baker *et al.*[18]
1987	*Int3* was discovered.	Gallahan and Callahan[19]
1988	Expression of the *int1* transgene under MMTV transcriptional regulator led to cancer in mice.	Tsukamoto *et al.*[20]
1989	*β*-Catenin was found to interact with E-cadherin.	Ozawa *et al.*[21]
1989	*Int1* is an organizer and embryonic inducer.	McMahon and Moon[22]
1990	*Int4* was discovered.	Roelink *et al.*[23]
1990	A large family of genes of *int1* was identified; *int* gene nomenclature became inadequate.	Gavin *et al.*[24]
1990	*β*-Catenin was characterized.	Takeichi[25]
1990	Adenomatous Polyposis Coli (APC) was discovered.	Groden *et al.* [26] Kinzler *et al.*[27]
1990	APC and *β*-catenin were involved in regulating adhesion between cells.	Peifer and Wieschaus[28]
1991	Wnt was given as a new name to *int1*, *wingless*, and their related genes.	Nusse *et al.*[29]
1992	*Int1* mutation phenomena are the consequences of *Wnt* gene mutations.	Nusse and Varmus[30]
1992	*Wnt*/*Wingless* is a Glycogen Synthase Kinase 3 (GSK3) inhibitor.	Siegfried *et al.* [31]
1993	APC was found interacting with *β*-catenin.	Rubinfeld *et al.*[32]
1993	*β*-Catenin is involved in Wnt/Wingless signaling.	Peifer *et al.*[33]
1994	Increased level of stabilized *β*-catenin was found in Wnt-activated cells.	Peifer *et al.*[34] Van Leeuwen *et al.*[35]

By 1995, an outline of the Wnt signaling pathway was generated, but there were still missing pieces in this network (**Figure 1.1**). In the 1990s, as more components were found on both upstream and downstream sides of the pathway, this signaling map started to connect (**Table 1.2**, **Figure 1.1**). Amongst the interactions, the discoveries of the *β*-catenin/T-cell factor (Tcf) interactions closed gaps between the Wnt signaling pathway and cancers.[3] Numerous genes are now known to have Tcf binding sites and therefore become transcriptional targets for the Wnt signaling pathway, and some of these genes such as *c-myc* are known to cause cancer.[36]

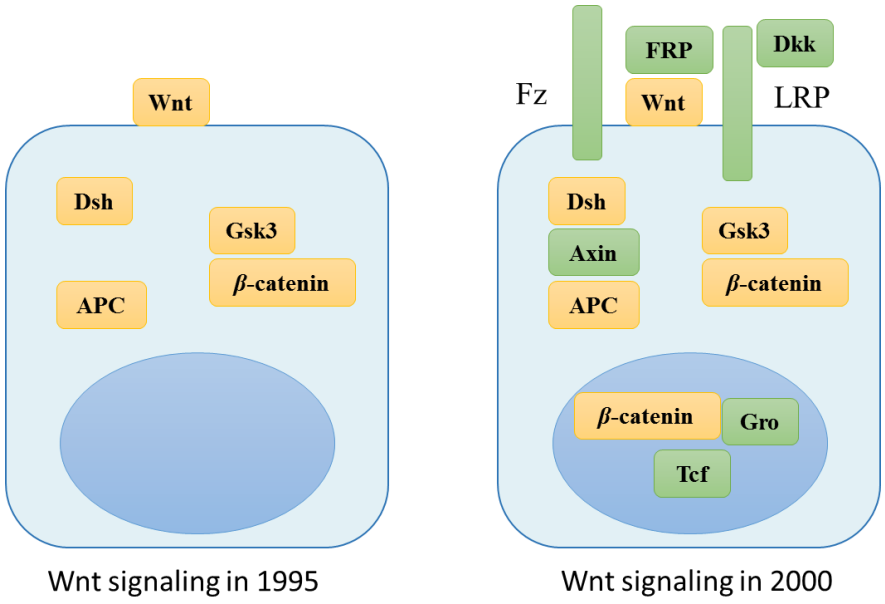


Figure 1.1. Wnt signaling in 1995 and 2000. Picture adapted from reference 3.

Table 1.2. Wnt components discovered after 1995.

Year	Event	Authors
1995	*β*-Catenin was found in cell nucleus and cell membrane	Funayama *et al.*[37]
1996	GSK3 was found in a complex with *β*-catenin and APC	Rubinfeld *et al.*[38]
1996	Tcf/Lef1 could interact with *β*-catenin	Molenaar *et al.*[39]
1996	Wingless protein can bind to Frizzled; transfection of Frizzled genes activates Wnt signaling	Bhanot *et al.*[40]
1997	FRP contains Wnt receptor domains and can bind to Wnt proteins	Finch *et al.*[41] Leyns *et al.*[42] Rattner *et al.*[43]
1997	Wingless signaling is mediated by TCF/Lef1	Riese *et al.*[44]
1998	Axin together with APC and GSK3 participates in the *β*-catenin destruction complex	Behrens *et al*[45]. Ikeda *et al.*[46]
2000	Lipoprotein receptor-related protein (LRP) is a co-receptor for Wnt, adjacent to Frizzled in the cell membrane	Pinson *et al.*[47] Tamai *et al.*[48]
2001	LRP-6 is a receptor for Dickkopf proteins	Mao *et al.*[49]

1.2 Cancer and Wnt Pathway

The canonical Wnt signaling pathway has been proven to play a crucial role in cell proliferation, differentiation, and survival.[50] The role of *Wnt1* in cancer was described in the mouse model of mammary and colon cancer as early as 1982.[10, 51] The direct connection between the Wnt pathway and human disease was first reported in the 1990s. The signaling pathway is regulated by Wnt ligands such as the APC-Axin complex and *β*-catenin. In colorectal and other cancers, these complexes are often found disrupted,[52-56] resulting in the stabilization of *β*-catenin, a multifunctional protein and a key mediator of the canonical Wnt pathway (**Figure 1.2**).

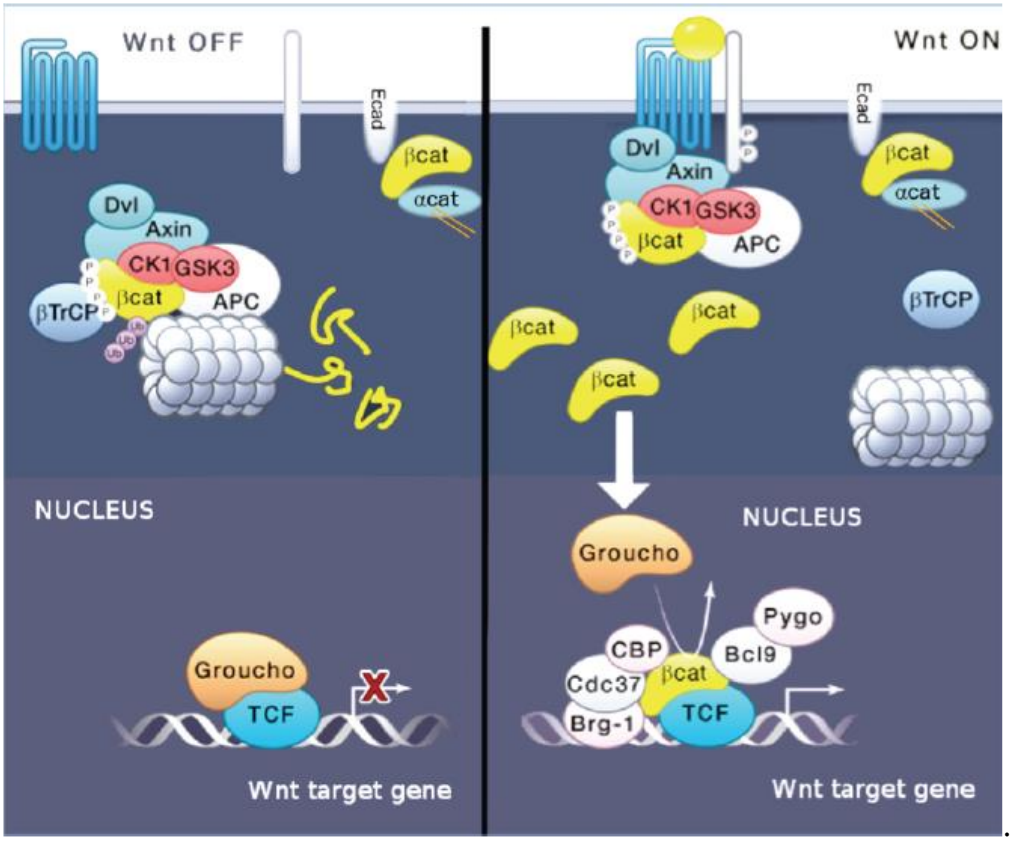


Figure 1.2. Wnt signaling pathway.[50] Picture adapted from reference 50.

In the absence of a Wnt signal, β-catenin is degraded by the destruction complex in the cytoplasm. APC and axin together facilitate the phosphorylation of β-catenin by casein kinase 1α (CK1α) and glycogen synthase kinase 3β (GSK3β) at regularly spaced *N*-terminal Ser/Thr residues. Phosphorylated β-catenin is then ubiquitinated by β-transducin repeats containing proteins (β-TRCP) leading to proteasomal degradation, maintaining a low level of free β-catenin to bind with Tcf proteins. Hence Tcf can interact with co-repressors such as Groucho and prevent target gene expression.

However, when the Wnt signaling is turned on, fizzled/low-density lipoprotein receptor-related protein (Fz/LRP) receptors are engaged by Wnt proteins, and Axin is recruited to the phosphorylated tail of LRP. This interaction inhibits the ubiquitination of β-catenin and therefore inhibits the degradation of phosphorylated β-catenin. As newly synthesized β-catenin cannot be degraded, it accumulates in the cell and translocates to the nucleus to form a complex with Tcf. The β-catenin/Tcf complex recruits transcriptional co-activators (CBP, Brg-1, Cdc73, and BCL9) and results in the expression of Wnt target genes, such as *c-myc* and *cyclin D1*.[50, 57-58] It was found that both of these genes contribute to the development of cancers.[36, 59-62] Past research also showed that β-catenin-Tcf complex (present in all of the nuclei of Wnt-related types of cancer) leads to chronic activation of a genetic program promoting cancer formation by stimulating cell growth and blocking apoptosis.[63] Disruption of the complex formation in colon cancer cells was proven to block the target gene's activation and to inhibit their *in vitro* growth.[61-62] Therefore, it is expected that drugs designed to disrupt this β-catenin/Tcf interaction will have great potential for treating Wnt-related cancers.[57]

1.3 Wnt Singling Pathway as Drug Target

Ever since the late 1990s, there has been significant interest in the development of effective Wnt pathway inhibitors in the pharmaceutical and biotechnology fields.[53-55] Methods such as high-throughput screening (HTS) and rational drug design have already identified some inhibitors, after nonsteroidal anti-inflammatory drugs showed no promising application in the chemoprevention or treatment of Wnt-related cancers.[57, 64-65]

Inhibitors of the upstream sites are less desirable because they can cause cross-regulatory effects on the β-catenin-independent Wnt pathways or perturb the function of β-catenin in cell-cell adhesion. Therefore, it would be advantageous to identify inhibitors for the downstream sites of the canonical Wnt signaling pathway. The formation of the β-catenin/Tcf complex in the cell nucleus is the penultimate step of canonical Wnt signaling. Transcriptional overactivation of canonical Wnt target genes is solely dependent on the formation of this complex. The deletion of the β-catenin gene can significantly reduce the growth of cancer cells.[66] Overexpression of dominant-negative Tcf,[62] or the use of siRNAs[67] and inducible shRNAs[68] against β-catenin can markedly decrease β-catenin-dependent gene expression and diminish the growth of many cancer cells *in vitro* and *in vivo*. Also, β-catenin knockdown by inducible shRNA abolishes cloning efficiency and the *in vivo* leukemogenic potential of MLL Leukemic stem cells.[69] Therefore, selective inhibition of β-catenin/Tcf protein-protein interactions (PPIs) is an appealing therapeutic target. Moreover, it is important that the inhibitor binds to co-activator β-catenin rather than transcriptional factor Tcf, because Tcf is essential for the other signaling pathways, while β-catenin is specific for canonical Wnt signaling.[70]

It is known that protein-protein interfaces tend to be flat, featureless and large

(typically 1500−3000 Å2),[71-72] and the surfaces of protein often display complex dynamical behavior, all of which greatly complicate the process of inhibitor design. In addition, the dissociation constant (K_d) of the β-catenin/Tcf complex is 7–10 nM[73-74] and much lower than those of the β-catenin/cadherin (K_d: 41–82 nM)[75] and β-catenin/APC (K_d: 0.6–3.1 μM for APC-R3, the APC repeat with the highest binding affinity) complexes,[76-77] which further complicates the design of selective β-catenin/Tcf inhibitors, making the disruption of β-catenin/Tcf interactions a "mission impossible".[57, 78]

1.4 Previously Identified Small Molecule Inhibitors

Lepourcelet *et al*. reported six small-molecule inhibitors of β-catenin/Tcf PPIs through high-throughput screening in 2004.[2] Surface plasmon resonance (SPR) assay and enzyme-linked immunosorbent assay (ELISA) were used in their primary screening to filter out six compounds from 7000 natural compounds. These compounds can all inhibit β-catenin/Tcf interaction with a low micromolar IC_{50} (**Figure 1.3**).

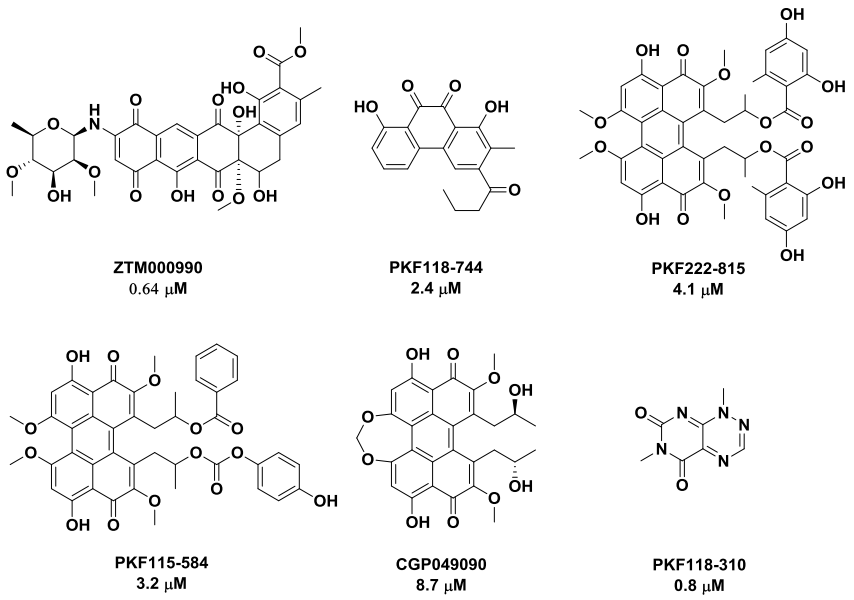


Figure 1.3. Inhibitors identified by HTS. IC_{50} values were measured in ELISA assay.

These compounds behave differently in their ability to inhibit Wnt signaling pathway *in vivo*. The three inhibitors from fungal derivatives PKF115–584, PKF118-310, and CGP049090 exhibit exceptional potency in inhibiting *β*-catenin-regulated transcription (CRT), colon cancer cell proliferation and *β*-catenin mediated axis duplication in Xenopus embryos. The exact mechanisms of how these inhibitors interfere with the target interaction were unclear. However, strong interference with *β*-catenin/APC interactions were observed for these compounds, suggesting that *β*-catenin was likely the target of the interaction.

In 2006, Trosset *et al.* reported a small molecule inhibitor PNU74654, identified through virtual screening and medium-throughput biophysical (NMR and ITC) assays (**Figure 1.4**).[79] The Isothermal Titration Calorimetry (ITC) experiments confirmed that the small molecule binds to *β*-catenin with a K_d of 450 nM. They hypothesized that the compound is likely to bind to the K435/R469 "hot spot" area on the *β*-catenin. However, no actual experimental support was given to verify the binding mode.

Tian *et al.* reported an organocopper inhibitor, BC21 (**Figure 1.5**), from virtual screening and luciferase reporter assays.[80] Its inhibitory activity against *β*-catenin/Tcf interactions was confirmed by the fluorescence polarization (FP) assay. BC21 also exhibits direct interruption of *β*-catenin/Tcf4 interaction and down-regulates the Wnt target genes, resulting in cancer cell death. They proposed that the binding site for BC21 was the pocket around K435 but had no experimental evidence to support this assignment.

More recently, Gonsalves *et al.* identified three small molecule inhibitors (iCRT3, iCRT5, and iCRT14) out of 14,977 compounds using high-throughput screening (**Figure 1.6**).[81] The HTS methods employed an Axin-specific dsRNA to address the selectivity for *β*-catenin/Tcf interaction over *β*-catenin/E-cadherin and other proteins. These compounds

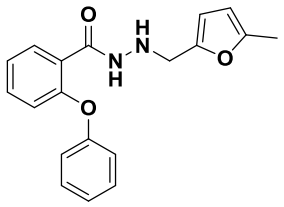


Figure 1.4. The structure of PNU74654. The K_d value was determined by ITC.

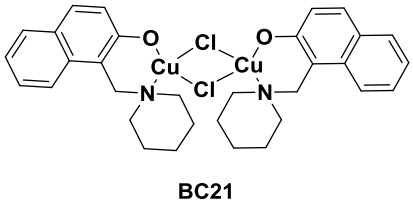


Figure 1.5. The structure of BC21.

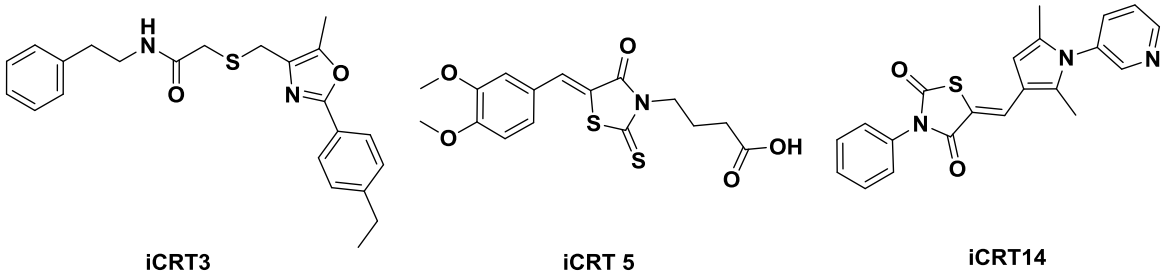


Figure 1.6. Structures of iCRT compounds.

inhibited the expression of Wnt target genes and the growth of colorectal cancers *in vitro* and *in vivo* through direct inhibition of *β*-catenin/Tcf4 interactions. Although *in silico* studies suggested over 20 druggable pockets on the *β*-catenin surface, the pocket lined by K435 and R469 was considered the most important binding site for the iCRT compounds, because these two residues are crucial for stabilizing the interaction for *β*-catenin/Tcf4.[82]

The druggability of these hit compounds was questionable as the conjugated quinone substructure in PKF115–584 and CGP049090, and the rhodanine substructure in iCRT5 and iCRT14 have been identified as frequent hits.[83] Moreover, the exact binding modes of these HTS hits remain unknown and hence have prohibited their optimization. Additionally, the selectivity of these inhibitors for *β*-catenin/Tcf over *β*-catenin/E-cadherin and *β*-catenin/APC interactions was low,[84] thus lowering the likelihood of therapeutic use.

Grossmann *et al.* reported that a hydrocarbon-stapled Axin *α*-helix, aStAx-35, can inhibit *β*-catenin/Tcf interactions by binding with the Axin-binding site of *β*-catenin.[85] However, as Axin is the scaffolding protein for *β*-catenin phosphorylation, the *β*-catenin/Axin interactions maintain a low level of free *β*-catenin in the cytoplasm. The use of these hydrocarbon-stapled peptides has the risk of elevating the concentrations of free *β*-catenin in normal cells, potentially promoting more *β*-catenin entering the cell nucleus and binding with Tcf. Therefore, it is important to find inhibitors that specifically target the *β*-catenin/Tcf interaction and leave the other two interactions unaffected.

1.5 Ji Laboratory Assay Development

In the Ji laboratory, efforts have been made to design and synthesize small molecule inhibitors of *β*-catenin/Tcf interactions. Homogeneous high-throughput fluorescence

polarization (FP) and AlphaScreen assays were first established for inhibitors of *β*-catenin/Tcf PPIs.[86]

The FP assay (**Figure 1.7**) used C-terminal fluorescein-labeled human Tcf4 (residues 7-51) and N-terminal His6-tagged human *β*-catenin (residues 138-686).[86] When the fluorescence tracer binds to the target protein, it will emit a strong FP signal. As the inhibitors disrupt the PPIs, the FP signal that is emitted will have a low polarization.

AlphaScreen is a bead-based technique. It is similar to the time-resolved fluorescence resonance energy transfer (TR-FRET) technique and can provide orthogonal results with respect to the FP assay. In the AlphaScreen assay, the streptavidin-coated donor beads and nickel-chelate acceptor beads are brought together through C-terminus biotinylated human Tcf4 (residue 7−51) and N-terminus His6-tagged human *β*-catenin (residues 138−686) interaction (**Figure 1.8**). Upon excitation by 680 nm laser light, the donor beads convert ambient oxygen to a singlet state. The lifetime of singlet oxygen is only a few microseconds,[87] and can only get to the acceptor beads if it is within 200 nm. If the PPIs are disrupted, the singlet oxygen can no longer reach the acceptor beads. Therefore no signal can be observed.

These two assays can exclusively evaluate the inhibitors ability to disrupt the protein-protein interaction complex. Previously identified inhibitors were re-evaluated using the FP assay. The results are shown in **Table 1.3**. The IC_{50}'s of the known inhibitors are reported as discovered by their source articles.

PKF115-584, CGP049090, PKF118-310, and PNU74654 were purchased through commercial sources, and the iCRT compounds were synthesized in the Ji lab. The synthetic routes are shown in **Figures 1.9—1.11**.

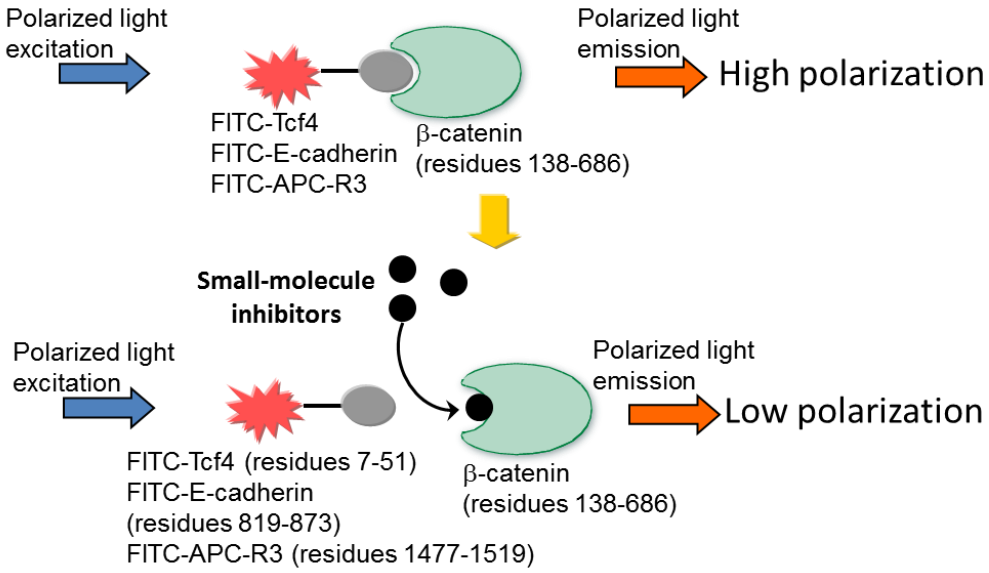


Figure 1.7. Schematic illustration of the FP assay reported by Zhang *et al.*.[86]

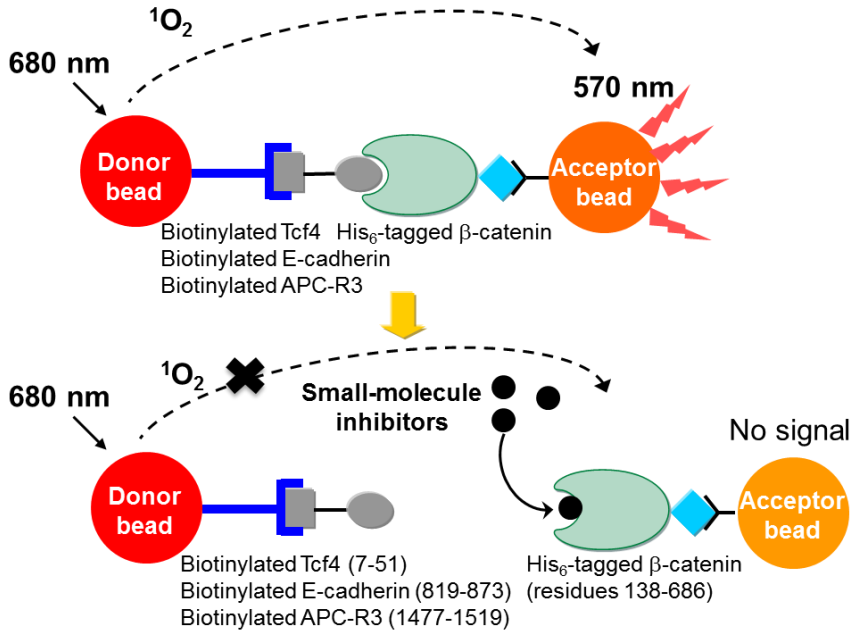


Figure 1.8. Schematic illustration of the AlphaScreen assay reported by Zhang *et al.*.[86]

Table 1.3. The binding affinity (K_i) of the known inhibitors measured in the Ji lab.[88]

Compound	K_i ± SD (μM) FP assay	IC_{50} (nM) reported
PKF115-584	17.78 ± 2.13	3200
CGP049090	35.56 ± 4.43	8700
PKF118-310	5.82 ± 0.24	800
PNU74654	180.99 ± 0.68	ND
iCRT3	364.71 ± 6.32	8.2
iCRT5	80.34 ± 4.17	18
iCRT14	53.51 ± 5.24	40.3

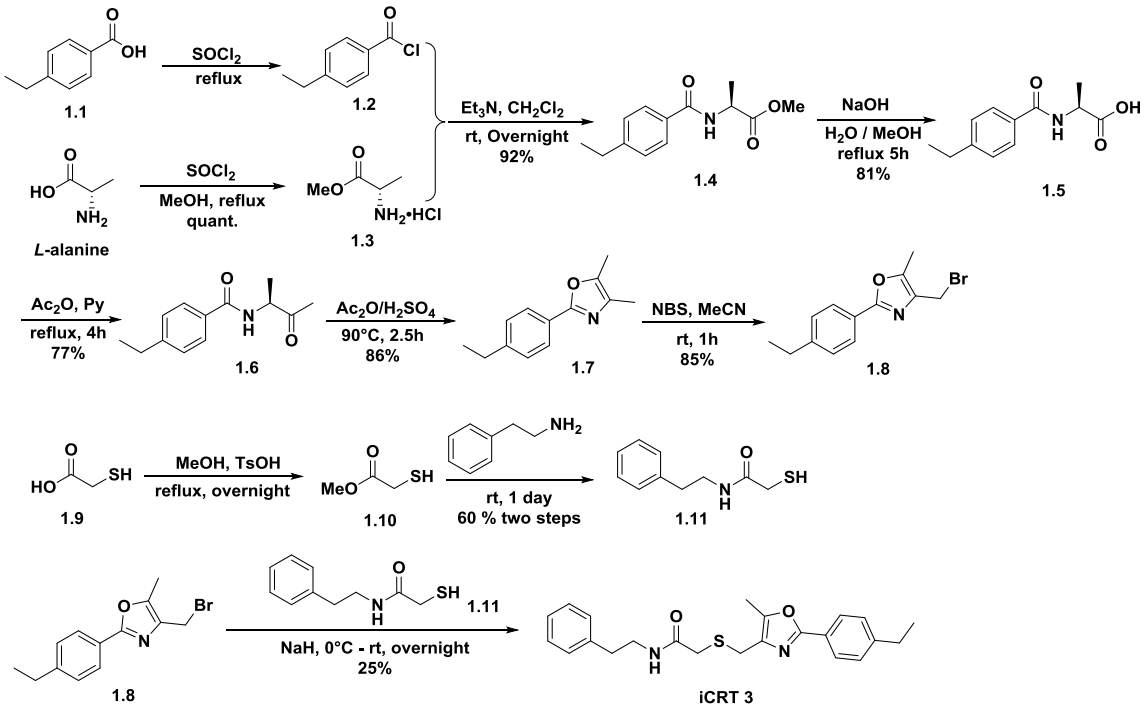


Figure 1.9. Synthesis of iCRT3.

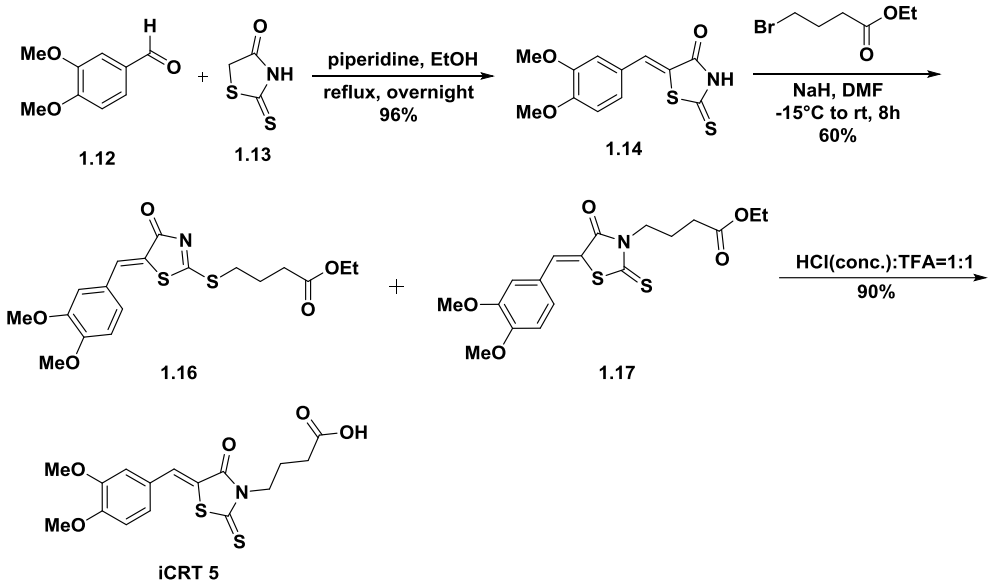


Figure 1.10. Synthesis of iCRT5.

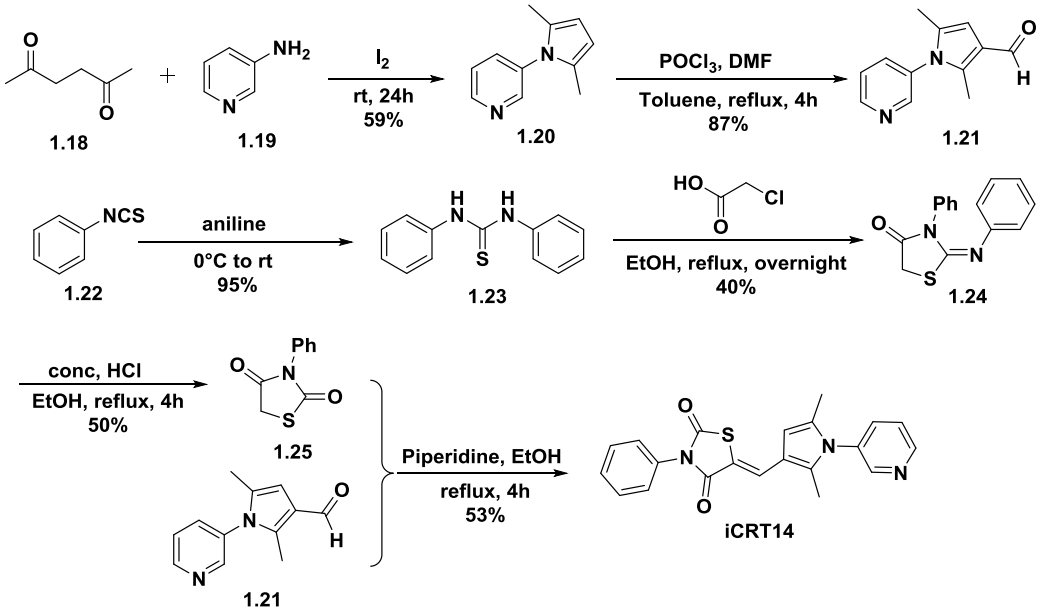


Figure 1.11. Synthesis of iCRT14.

Synthesis of iCRT3. Key intermediate 4,5-dimethyloxazole (**1.6)** was synthesized using a Dakin-West reaction in excellent yields.[89-90] Selective halogenation on the 4-methyl group was achieved by treating **1.6** with one equivalent of NBS in acetonitrile at room temperature.[91] Treating **1.11** with NaH at 0˚C and brominated oxazole **1.8** gave the iCRT3 in 10% overall yield.

Synthesis of iCRT5. The Knoevenagel condensation[92] between vanillin aldehyde **1.12** and rhodanine **1.13** was carried out using a reported method to afford the adduct **1.14**.[93] Alkylation on the sulfur (**1.16**) is favored over nitrogen. Lower temperature together with a hard counter ion (Na^+) yielded slightly more of the desired *N*-alkylated product **1.17**. Attempts to hydrolyze the ethyl ester in **1.17** under basic conditions resulted in a complete decomposition of the rhodanine. In contrast, **1.17** was very tolerant of harsh acidic condition (12 M aqueous HCl and trifluoroacetic acid), yielding iCRT5 in 46% overall yield.

Synthesis of iCRT14. The two key intermediates for the synthesis of iCRT14 pyrrole-3-carbaldehyde (**1.21**) and 3-phenylthiazolidine-2,4-dione (**1.25**) were synthesized in satisfactory yields according to reported methods.[94-95] Exposure of **1.21** and **1.25** to Knoevenagel condensations condition generated iCRT14 with a 10% overall yield.

1.6 "Hot Spot" Based Design

It was reported that not all interfaces between the two proteins are equally important for PPIs.[96] Small subsets of residues on protein surface contribute to most of the free energy of binding,[97] termed as "hot spots," which can serve as the starting point for inhibitor design. Although Tcf and *β*-catenin have a large protein-protein contacting

surface (≥2800 Å^2 versus typical 1500-3000 Å^2), three hot spots on β-catenin for binding to Tcf were revealed by crystallographic and biochemical analyses,[66, 82, 98-100] shown in **Figure 1.12**.[88] Hot spot 1 (K435/K508 binding to D16/E17 of Tcf4) and 2 (K312/K345 binding E24/E29 of Tcf4)[99, 101] are charge-charge interactions, while the third one is a hydrophobic pocket that consists of F253, I256, I293, A295 and I296, interacting with the hydrophobic residues (V44 and L48) on Tcf4.

It has been evident that K435 is the most important "hot spot" for T-cell factor (Tcf) binding.[74, 98, 100, 102-103] The D16A mutation of human Tcf4 results in a complete loss of its ability to bind to β-catenin.[100, 102] Additionally, the D19A or E20A mutation of mouse Lef-1 (corresponding to D16 and E17 of human Tcf4, respectively) prevents nuclear translocation of β-catenin,[82] and the simultaneous mutation of D19A and E20A results in a protein unable to bind to β-catenin.

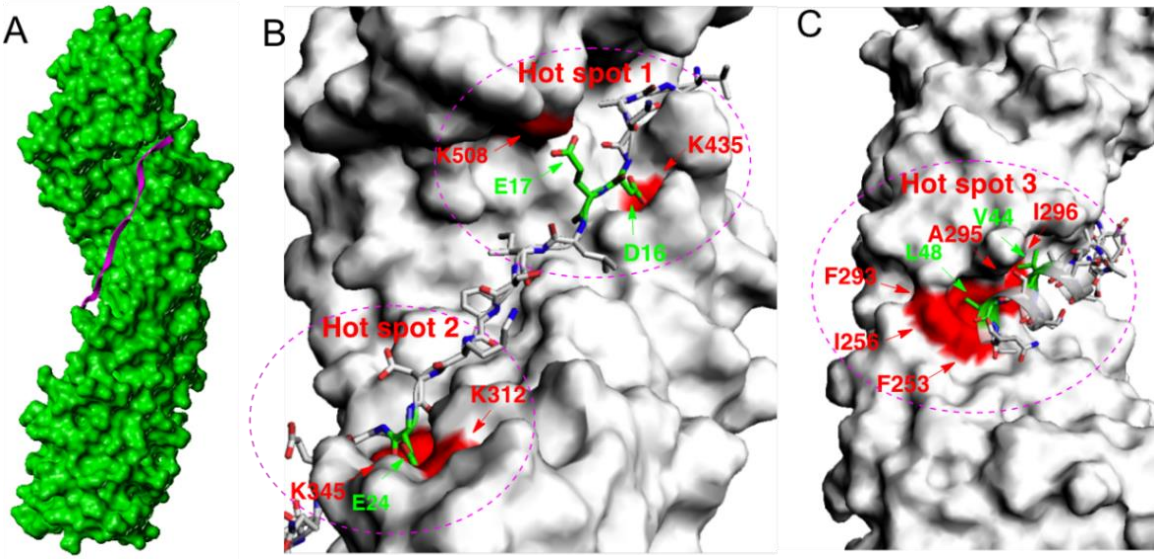


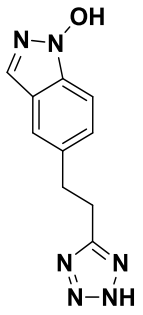
Figure 1.12.[88] Structure of the β-catenin/Tcf complex (PDB id: 2GL7[78]).
A. Armadillo repeats 5-9 of β-catenin binds to Tcf. β-Catenin is shown as a surface model (green), and Tcf is shown as a ribbon (purple).
B. "Hot spots" 1 and 2. β-Catenin: surface model (gray), and Tcf: stick model. "Hot spot" 1: K435 and K508 of β-catenin. "Hot spot" 2: K312 and K345 of β-catenin.
C. "Hot spot" 3: F253, I256, F293, A295 and I296 of β-catenin.

Yu *et al.* used alanine scanning combined with SPR experiments to evaluate the contribution of each hot spot (**Table 1.4**). The studies confirmed that hot spot 1 (D16/E17 on Tcf4) was the most important, as Tcf4 completely lost the ability to bind to *β*-catenin after the mutation.

The SPR study strongly indicated that mimicking D16/E17 was a rational starting point and feasible approach to designing small molecule inhibitors for *β*-catenin/Tcf PPIs. Therefore, a small–molecule inhibitor **UU-01** (**Figure 1.13**) was designed and synthesized for the *β*-catenin/Tcf4 interaction by mimicking the D16/E17 residues of Tcf4 and bioisosteric replacement strategy.[88] The K_i was determined by competitive FP assay to be 3.14 ± 0.48 μM.[86, 88]

Table 1.4. SPR study of *β*-catenin with wild-type and mutant Tcf4 peptides.[88]

Entry	Tcf4 peptides	K_d (nM)
1	Residues 7-15	4.71
2	Residues 7-15, D16A, E17A	>100,000
3	Residues 7-15, E24A, E29A	28
4	Residues 7-15, V44A, L48A	89.3



UU-01

Figure 1.13. Small-molecule inhibitor reported by Yu *et al.*

However, though **UU-01** showed excellent inhibitory activity towards *β*-catenin/Tcf4 interaction, its selectivity towards *β*-catenin/APC and *β*-catenin/E-cadherin interactions was not carried out. Designing a selective inhibitor of this target remained a challenge.

1.7 Conclusion

The Wnt signaling pathway has become an important target for designing drugs for cancer therapy. This work focused on the design of a selective inhibitor for *β*-catenin/Tcf4 PPIs using a peptidomimetic strategy. The peptidomimetic strategy faces challenges when designing small-molecule PPIs inhibitors due to the large protein−protein surface and the noncontiguous distribution of hot spots.[71] Peptides derived from short contiguous sequences at the protein−protein interface have been shown to be poor chemical starting points for disrupting PPIs. The current peptidomimetic strategies, however, strive to mimic protein hot spot residues[88, 104] or protein secondary structures,[105] including *α*-helix, *β*-sheet, and *β*-turn. In this work, we hypothesized that a peptidomimetic strategy could be used to generate new potent inhibitors through the exploration of new pockets/new binding elements that are adjacent to a hot spot but not employed by *β*-catenin/Tcf4 PPIs.

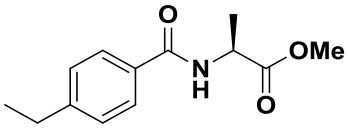
1.8 Methods and Experiments

1.8.1 Chemistry

All experiments were conducted under anhydrous conditions under an argon atmosphere, using flame-dried apparatus and employing standard techniques in handling air-sensitive materials. Dichloromethane (CH_2Cl_2), acetonitrile (CH_3CN), tetrahydrofuran

(THF), dimethylformamide (DMF) were degassed with nitrogen and passed through JC Meyer solvent systems. All reagents were purchased from Sigma-Aldrich and used as received. Aqueous solutions of sodium bicarbonate ($NaHCO_3$), sodium chloride (brine), and ammonium chloride (NH_4Cl) were saturated. Analytical thin layer chromatography was visualized by ultraviolet light. Flash chromatography was performed on SilicaFlash@ F60 silica gel (230–400 mesh). ^1H NMR spectra were recorded using a Varian Unity Inova 500 (500 MHz) or a Varian Unity Inova 300 (300 MHz). ^{13}C NMR spectra were recorded using a Varian Unity Inova 500 (125 MHz) or Varian Unity Inova 300 (75 MHz). The ^1H and ^{13}C NMR spectra were referenced to the residual solvent signals (7.26 ppm for ^1H and 77.0 ppm for ^{13}C in $CDCl_3$; 2.50 ppm for ^1H and 39.5 ppm for ^{13}C in DMSO-d6). All melting points were determined on a Büchi M560 melting point apparatus and uncorrected. Low (MS) and high (HRMS) mass spectra were determined on a Micromass Quattro II (ESI/APCI-TOF) at the University of Utah, Department of Chemistry Mass Spectrometry Facility.

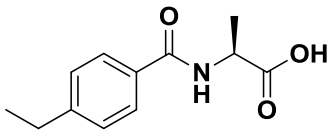
1.8.2 Synthetic Routes, Procedures and Characterizations



1.4

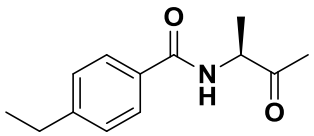
(*S*)-Methyl 2-(4-ethylbenzamido) propanoate (1.4). To a stirred solution of 4-ethylbenzoyl chloride (0.786 g, 4.7 mmol) in 10 mL dichloromethane was added (*S*)-methyl 2-aminopropanoate hydrochloride (0.700 g, 5 mmol) and 1.4 mL Et_3N (10 mmol). The reaction was stirred at rt overnight. The mixture was poured into 20 mL 1 M HCl (aq), and extracted with ethyl acetate (2 × 15 mL). The organic layers were combined and

washed with saturated $NaHCO_3$ (aq) (2 × 20 mL) and brine (2 × 20 mL), then dried and evaporated to afford 0.968 g product as a white solid (92%). ^{1}H NMR (300 MHz, Chloroform-*d*) δ 7.75 (m, 2H), 7.28 (m, 2H), 6.72 (d, *J* = 6.4 Hz, 1H), 4.82 (p, *J* = 7.2 Hz, 1H), 3.81 (s, 3H), 2.71 (q, *J* = 7.6 Hz, 2H), 1.54 (d, *J* = 7.1Hz, 3H), 1.26 (t, *J* = 7.6 Hz, 3H). ^{13}C NMR (125MHz, $CDCl_3$) δ 174.0, 166.9, 148.7, 131.5, 128.3, 127.4, 52.8, 48.6, 29.0, 18.9, 15.5. MS (ESI) *m/z* = 236.4 $[M + H]^+$.



1.5

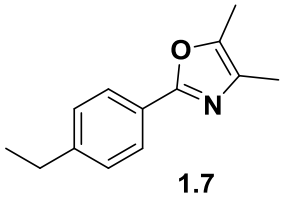
(*S*)-2-(4-Ethylbenzamido) propanoic acid (1.5). To a stirred solution of **1.4** (0.968 g, 4.34 mmol) in 5 mL of MeOH was added 5 mL 1M NaOH. The reaction was heated to reflux for 5 h. The reaction mixture was poured into 15 mL of 1M NaOH and washed with dichloromethane (2 × 10 mL). Then the aqueous layer was acidified with 1 M HCl (aq) to pH = 3, and extracted with ethyl acetate (3 × 15 mL). The organic layer was combined and dried then concentrated *in vacuo* to yield pure product as white solid (0.797 g, 81%). ^{1}H NMR (300 MHz, Chloroform-*d*) δ 7.78 – 7.67 (m, 2H), 7.32 – 7.22 (m, 2H), 6.70 (d, *J* = 6.9 Hz, 1H), 4.79 (p, *J* = 7.1 Hz, 1H), 2.70 (q, *J* = 7.6 Hz, 2H), 1.58 (d, *J* = 7.2 Hz, 3H), 1.24 (t, *J* = 7.6 Hz, 3H). ^{13}C NMR (75 MHz, $CDCl_3$) δ 176.1, 149.2, 130.8, 128.4, 127.4, 49.0, 29.0, 18.3, 15.5. MS (ESI) *m/z* = 222.6 $[M + H]^+$.



1.6

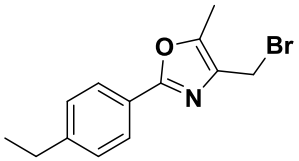
4-Ethyl-*N*-(3-oxobutan-2-yl) benzamide (1.6). To a stirred solution of **1.5** (0.72

g, 3.25mmol) in 7.5 mL pyridine was added 0.72 mL acetic anhydride. The reaction was heated to reflux for 2.5 h. The mixture was dissolved in 30 mL of ethyl acetate and washed with 1M HCl (aq) (2 × 20 mL), 10% Na_2CO_3 (aq) (2 × 20 mL) and brine (2 × 20 mL), then dried and purified by silica gel column chromatography (hexanes : ethyl acetate = 1 : 1) to afford 0.552 g product as a white solid (77%). ^{1}H NMR (300 MHz, Chloroform-*d*) δ 7.79 – 7.70 (m, 2H), 7.33 – 7.22 (m, 2H), 7.02 (d, *J* = 6.4 Hz, 1H), 4.81 (p, *J* = 7.0 Hz, 1H), 2.71 (q, *J* = 7.6 Hz, 2H), 2.30 (s, 3H), 1.50 (d, *J* = 7.2 Hz, 3H), 1.26 (t, *J* = 7.6 Hz, 3H). ^{13}C NMR (75 MHz, $CDCl_3$) δ 207.2, 148.7, 131.6, 128.3, 127.3, 55.2, 29.0, 26.9, 18.0, 15.6. MS (ESI) *m/z* = 220.5 $[M + H]^+$.



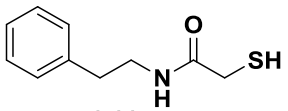
1.7

2-(4-Ethylphenyl)-4,5-dimethyloxazole (1.7). To a stirred solution of **1.6** (0.552 g, 2.5 mmol) in 1.5 mL acetic anhydride was added 1 mL concentrated H_2SO_4. The reaction was heated to reflux for 2.5 h. Upon completion, the mixture was poured into 30 mL 1M NaOH and extracted with ethyl acetate (2 × 20 mL). The organic layers were combined and washed with 10% Na_2CO_3 (aq) (2 × 30 mL) and brine (2 × 30 mL) dried. The solvent evaporated *in vacuo* to afford 0.432 g product as a white solid (86%). ^{1}H NMR (300 MHz, Chloroform-*d*) δ 7.94 – 7.80 (m, 2H), 7.29 – 7.20 (m, 2H), 2.68 (q, *J* = 7.6 Hz, 2H), 2.30 (d, *J* = 0.9 Hz, 3H), 2.15 (q, *J* = 0.9 Hz, 3H), 1.25 (t, *J* = 7.6 Hz, 3H). ^{13}C NMR (75 MHz, $CDCl_3$) δ 159.6, 146.3, 143.2, 131.9, 128.4, 126.1, 125.6, 29.0, 15.6, 11.5, 10.3. MS (ESI) *m/z* = 202.6 $[M + H]^+$.



1.8

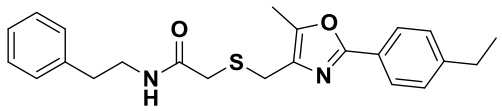
4-(Bromomethyl)-2-(4-ethylphenyl)-5-methyloxazole (1.8). To a stirred solution of **1.7** (0.419 g, 2.08 mmol) in 4 mL MeCN at rt was added NBS (0.407 g, 2.29 mmol). The reaction was stirred at rt for 1.5 h. Then the reaction mixture was poured into 20 mL 1M HCl (aq), and then extracted with ethyl acetate (2 × 15 mL). The organic layers were combined and washed with 10% Na_2CO_3 (aq) (2 × 30 mL) and brine (30 mL), then concentrated under reduced pressure. The residue was purified by chromatography (hexanes : ethyl acetate = 4 : 1) to afford 0.546 g product as a white solid (0.546 g, 85%). ^{1}H NMR (300 MHz, Chloroform-*d*) δ 7.98 – 7.86 (m, 2H), 7.31 – 7.08 (m, 2H), 4.44 (s, 2H), 2.69 (q, *J* = 7.6 Hz, 2H), 2.39 (s, 3H), 1.26 (t, *J* = 7.6 Hz, 3H). ^{13}C NMR (75 MHz, $CDCl_3$) δ 159.8, 147.0, 128.5, 126.5, 124.9, 29.1, 24.2, 15.5, 10.7. MS (ESI) *m/z* = 280.5 $[M + H]^+$.



1.11

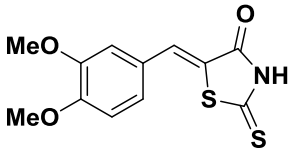
2-Mercapto-*N*-phenethylacetamide (1.11). To a stirred solution of 0.7 mL thioglycolic acid in 5 mL methanol was added 0.2 g 4-methylbenzenesulfonic acid. The reaction was heated to reflux overnight. After TLC showed no thioglycolic acid left, the reaction was cooled to rt and 1.26 mL 2-phenylethanamine was added. The reaction was stirred at rt overnight. Then the reaction was concentrated under reduced pressure and the residue was purified by column chromatography (hexanes: ethyl acetate = 1 : 4, R_f = 0.3) to afford 0.7 g product as a white solid. ^{1}H NMR (500 MHz, Chloroform-*d*) δ 7.31 (t, *J* =

7.8 Hz, 2H), 7.26 – 7.19 (m, 3H), 6.53 (s, 1H), 3.57 (q, *J* = 6.8 Hz, 2H), 3.33 (s, 2H), 2.87 (t, *J* = 7.0 Hz, 3H). ^{13}C NMR (75 MHz, $CDCl_3$) δ 168.5, 138.9, 129.0, 128.9, 126.8, 42.7, 41.4, 35.7. MS (ESI) *m/z* = 196.5 $[M + H]^+$.



iCRT 3

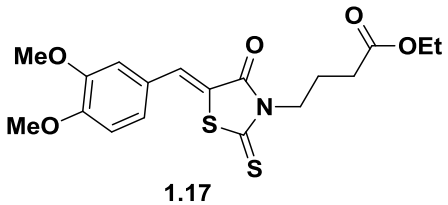
2-{{[2-(4-Ethylphenyl)-5-methyloxazol-4-yl] methyl}thio}-*N*-phenethylacetamide (iCRT3**).** To a stirred solution of **1.11** (0.12 g, 0.615 mmol) at -15 ˚C was added 0.03 g NaH (60%, 0.74 mmol) proportionally. After 30 min, **1.8** (0.21 g, 0.74 mmol) was added. The reaction was allowed to warm to rt slowly and stirred overnight. After the reaction was complete (monitored by TLC), the solvent was evaporated and the residue was purified by column chromatography (hexanes : ethyl acetate = 7 : 3 to 1 : 1) to afford 60 mg product as a pale yellow solid (25%). ^{1}H NMR (300 MHz, Chloroform-*d*) δ 7.87 (d, *J* = 8.4 Hz, 2H), 7.36 (s, 1H), 7.30 – 7.15 (m, 5H), 7.14 – 7.06 (m, 2H), 3.55 – 3.43 (m, 4H), 3.20 (s, 2H), 2.78 (t, *J* = 7.2 Hz, 2H), 2.70 (q, *J* = 7.5 Hz, 2H), 2.32 (s, 3H), 1.27 (t, *J* = 7.6 Hz, 3H). ^{13}C NMR (75 MHz, $CDCl_3$) δ 168.9, 160.3, 147.1, 145.3, 138.9, 132.3, 128.9, 128.8, 128.6, 126.7, 126.3, 124.9, 41.1, 35.7, 35.5, 29.1, 27.4, 15.6, 10.5. HRMS (ESI, CH_3OH) $C_{23}H_{27}N_2O_2S$ Calc.: 395.1793, Found: 395.1795.



1.14

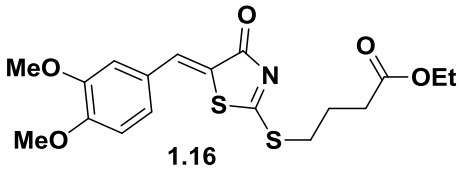
(*Z*)-5-(3,4-dimethoxybenzylidene)-2-thioxothiazolidin-4-one (**1.14**). To a stirred solution of 0.66 g Rhodanine (5 mmol) and 0.83 g methyl vanillin aldehyde (5 mmol) in 20 mL EtOH was added 0.15 mmol piperidine. The reaction was heated to reflux overnight.

Yellow solid was formed. The reaction was cooled in an ice bath. The solid was filtered and the filter cake was washed with 5 mL cold EtOH and 10 mL hexanes. The solid was dried in the oven for 30 min to afford the 1.396 g product as a yellow solid (99%).^{1}H NMR (500 MHz, DMSO-d_6) δ 7.59 (d, J = 3.6 Hz, 1H), 7.20 – 7.13 (m, 2H), 7.11 (d, J = 8.3 Hz, 1H), 3.82 (s, 3H), 3.80 (s, 3H). ^{13}C NMR (126 MHz, DMSO) δ 151.9, 149.7, 133.0, 126.3, 125.3, 123.0, 114.1, 112.9, 95.0, 56.4, 56.2. MS (ESI) *m/z* = 282.7 $[M + H]^+$.



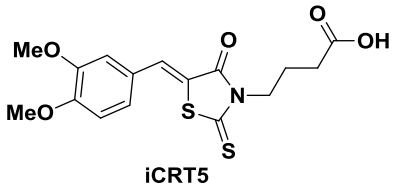
1.17

Ethyl (*Z*)-4-(5-(3,4-dimethoxybenzylidene)-4-oxo-2-thioxothiazolidin-3-yl)butanoate (**1.17**). At -10˚C, to a stirred solution of 1.379 g **1.14** (4.9 mmol) in 50 mL DMF was added 0.23 g NaH slowly (5.8 mmol). The reaction was stirred at -10˚C for 0.5 h and 0.84 mL ethyl 4-bromobutanoate (5.4 mmol) was added. The reaction was allowed to warm to the rt over 2 h and left to react overnight. Upon completion, the reaction mixture was poured into 100 mL 1 M HCl (aq) and extracted with ethyl acetate (50 mL × 2). The organic layers were combined and concentrated *in vacuo*. The residue was purified by column chromatography to afford 0.48 g **1.17** as a yellow solid (25%) and 0.35 g **1.16** as an orange sold (18%). ^{1}H NMR (500 MHz, Chloroform-*d*) δ 7.67 (s, 1H), 7.14 (dd, J = 8.4, 2.1 Hz, 1H), 7.03 – 6.93 (m, 2H), 4.20 (t, J = 7.1 Hz, 2H), 4.14 (q, J = 7.2 Hz, 2H), 3.95 (s, 3H), 3.95 (s, 3H), 2.39 (t, J = 7.4 Hz, 3H), 2.07 (p, J = 7.3 Hz, 3H), 1.26 (t, J = 7.1 Hz, 3H). ^{13}C NMR (125 MHz, $CDCl_3$) δ 193.5, 172.7, 168.1, 151.8, 149.8, 133.8, 126.5, 125.9, 120.3, 112.6, 111.7, 60.8, 56.3, 56.2, 43.9, 31.8, 22.7, 14.5. MS (ESI) *m/z* = 396.0 $[M + H]^+$.



1.16

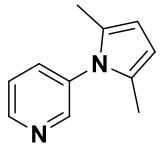
Ethyl (*Z*)-4-((5-(3,4-dimethoxybenzylidene)-4-oxo-4,5-dihydrothiazol-2-yl)thio)butanoate (**1.16**). ^{1}H NMR (300 MHz, Chloroform-*d*) δ 7.82 (s, 1H), 7.15 (ddd, *J* = 8.4, 2.2, 0.6 Hz, 2H), 7.01 (d, *J* = 2.1 Hz, 2H), 6.94 (d, *J* = 8.5 Hz, 1H), 4.14 (q, *J* = 7.1 Hz, 2H), 3.94 (d, *J* = 1.1 Hz, 9H), 3.53 – 3.41 (m, 2H), 2.47 (t, *J* = 7.2 Hz, 3H), 2.15 (p, *J* = 7.2 Hz, 3H), 1.26 (t, *J* = 7.1 Hz, 3H). ^{13}C NMR (75 MHz, $CDCl_3$) δ 172.7, 151.7, 149.6, 136.5, 126.7, 125.5, 123.8, 112.6, 111.6, 60.9, 56.3, 56.2, 33.0, 32.9, 24.7, 14.4.



iCRT5

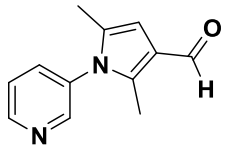
(*Z*)-4-(5-(3,4-dimethoxybenzylidene)-4-oxo-2-thioxothiazolidin-3-yl)butanoic acid (**iCRT5**). To a mixture of 3 mL concentrated HCl (aq) and 3 mL trifluoroacetic acid was added 100 mg 1.17. The reaction was heated to reflux overnight. Upon completion, the reaction was cool in an ice bath, the pH was adjusted to 4 by adding 2 M NaOH (aq). The mixture was then extracted with ethyl aceate (2 × 10 mL). The organic layers were combined and washed with brine (2 × 20 mL), dried and concentrated *in vacuo* to afford 92 mg product as a yellow solid. (99%). ^{1}H NMR (500 MHz, Chloroform-*d*) δ 8.59 (s, 2H), 7.66 (s, 1H), 7.12 (dt, *J* = 8.3, 2.0 Hz, 1H), 6.99 – 6.92 (m, 2H), 4.18 (t, *J* = 7.0 Hz, 2H), 3.94 (s, 2H), 3.93 (s, 2H), 2.44 (t, *J* = 7.3 Hz, 2H), 2.10 – 2.00 (m, 2H). ^{13}C NMR (126 MHz, $CDCl_3$) δ 193.4, 178.1, 168.4, 151.9, 149.7, 134.2, 126.4, 126.0, 120.0, 112.7, 111.7, 56.3, 56.2, 43.7, 31.2, 22.3. HRMS (ESI, CH_3OH) $C_{16}H_{17}NO_5S_2$ Calc.: 367.0548 Found:

390.0446 $[M+Na]^+$.



1.20

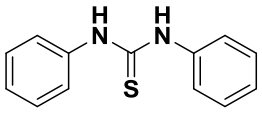
3-(2,5-Dimethyl-1*H*-pyrrol-1-yl)pyridine (**1.20**). To a stirred solution of 0.94 g 3-amino pyridine (10 mmol) in 1.4 mL 2,5-hexadione (12 mmol) at rt was added 0.25 g iodine crystals. The reaction was stirred at rt for 24 h. Upon completion, the reaction mixture was diluted in 50 mL ethyl acetate. The solution was washed with 50 mL 5% $Na_2S_2O_3$ (aq). The organic layer was then concentrated *in vacuo*. The residue was purified with column chromatography (hexanes : ethyl acetate = 8:2) to yield 1.02 g product as colorless oil (59%). ^{1}H NMR (300 MHz, Chloroform-*d*) δ 8.66 (dd, *J* = 4.9, 1.5 Hz, 1H), 8.54 (d, *J* = 2.5 Hz, 1H), 7.59 (dt, *J* = 8.1, 2.0 Hz, 1H), 7.44 (dd, *J* = 8.0, 4.8 Hz, 1H), 5.95 (s, 2H), 2.05 (s, 6H). ^{13}C NMR (75 MHz, $CDCl_3$) δ 149.6, 149.0, 135.7, 129.1, 123.9, 106.8, 13.2. MS (ESI) *m/z* = 172.9. $[M + H]^+$.



1.21

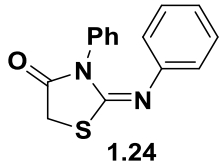
2,5-Dimethyl-1-(pyridin-3-yl)-1*H*-pyrrole-3-carbaldehyde (**1.21**). To a stirred solution of 1.01 g **1.20** (5.87 mmol) in 20 mL toluene was added 0.45 mL DMF (5.87 mmol) and 0.55 mL $POCl_3$ (5.87 mmol). The reaction was heated to reflux for 3.5 h. Upon completion, the reaction mixture was cooled to rt. Saturated NaOAc (aq) (20 mL) was added and stirred for 20 min. K_2CO_3 (2 g) was added to the mixture and the mixture was extracted with dichloromethane (2 × 30 mL). The organic layers were combined and concentrated *in vacuo*. The residue was purified by column chromatography to afford 1.02

g product as colorless oil (87%). ^{1}H NMR (300 MHz, Chloroform-*d*) δ 9.89 (s, 1H), 8.75 (dd, *J* = 4.7, 1.7 Hz, 1H), 8.53 (dd, *J* = 2.5, 0.8 Hz, 1H), 7.59 (ddd, *J* = 8.0, 2.5, 1.6 Hz, 1H), 7.50 (ddd, *J* = 8.1, 4.8, 0.8 Hz, 1H), 6.42 (q, *J* = 1.1 Hz, 1H), 2.29 (s, 3H), 2.00 (d, *J* = 1.0 Hz, 3H). ^{13}C NMR (126 MHz, $CDCl_3$) δ 174.0, 166.9, 148.7, 131.5, 128.3, 127.3, 52.8, 48.6, 29.0, 19.0, 15.5. MS (ESI) *m/z* = 201.2 $[M + H]^+$.



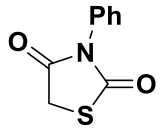
1.23

1,3-Diphenylthiourea (**1.23**). At 0˚C, 1 mL aniline (11 mmol) was added to 1.5 mL isothiocyanatobenzene dropwise with stirring. White solid was formed and the reaction became too thick to stir. The solids were recrystallized from hot ethanol to afford 2.23 g product as tabular crystal (89%). m.p. 152.2-154.5˚C. ^{1}H NMR (300 MHz, DMSO-d_6) δ 9.77 (s, 2H), 7.49 – 7.42 (m, 4H), 7.36 – 7.24 (m, 4H), 7.14 – 7.04 (m, 2H). ^{13}C NMR (75 MHz, DMSO) δ 180.2, 140.1, 129.1, 125.1, 124.3. MS (ESI) *m/z* = 229.4 $[M + H]^+$.



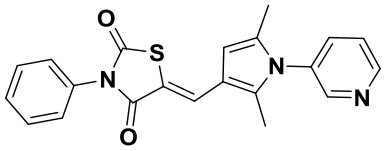
1.24

(*E*)-3-Phenyl-2-(phenylimino)thiazolidin-4-one (**1.24**). To a stirred solution of 1.12 g **1.23** (4.9 mmol) in 10 mL ethanol was added 0.46 g monochloro-acetic acid (4.9 mmol). The reaction as heated to reflux overnight. The reaction was then cooled in an ice bath, and white crystals were formed. The mixture was filtered and the solid was washed with cold ethanol to afford 0.527 g product as a white solid (40%). ^{1}H NMR (300 MHz, Chloroform-*d*) δ 7.53 (t, *J* = 7.5 Hz, 2H), 7.48 – 7.36 (m, 3H), 7.36 – 7.28 (m, 2H), 7.12 (t, *J* = 7.5 Hz, 1H), 6.92 (d, *J* = 7.8 Hz, 2H), 3.99 (s, 2H). MS (ESI) *m/z* = 269.5 $[M + H]^+$.



1.25

3-Phenylthiazolidine-2,4-dione (**1.25**). To a stirred solution of 0.527 g **1.24** (1.96 mmol) in 10 mL ethanol was added 1.5 mL concentrated HCl (aq). The reaction was heated to reflux for 5 h. The reaction was cooled in an ice bath and white precipitate formed. The mixture was filtered and the solid was washed with cold water and hexanes to afford 0.201 product as a white solid (53%). ^{1}H NMR (500 MHz, Chloroform-*d*) δ 7.54 – 7.46 (m, 2H), 7.49 – 7.38 (m, 1H), 7.29 – 7.22 (m, 2H), 4.12 (s, 2H). ^{13}C NMR (126 MHz, $CDCl_3$) δ 170.9, 132.9, 129.7, 129.6, 127.4, 34.1. MS (ESI) *m/z* = 193.9 $[M + H]^+$.



iCRT14

(*Z*)-5-((2,5-Dimethyl-1-(pyridin-3-yl)-1*H*-pyrrol-3-yl)methylene)-3-phenylthiazolidine-2,4-dione (**iCRT14**). To a stirred solution of 0.1 g **1.21** (0.5 mmol) and 0.101 g **1.25** (0.524 mmol) in 5 mL ethanol was added 2 drops of piperidine. The reaction was heated to reflux for 2 h. Upon completion, the reaction was cooled to rt and the solvent was concentrated *in vacuo*. The residue was purified by column chromatography to yield 0.103 g product as a white solid (55%). ^{1}H NMR (300 MHz, Chloroform-*d*) δ 8.73 (dd, *J* = 4.7, 1.6 Hz, 1H), 8.53 (dd, *J* = 2.5, 0.8 Hz, 1H), 7.95 (d, *J* = 0.5 Hz, 1H), 7.58 (ddd, *J* = 8.1, 2.5, 1.6 Hz, 1H), 7.55 – 7.38 (m, 4H), 7.38 – 7.29 (m, 2H), 6.33 – 6.28 (m, 1H), 2.18 (s, 3H), 2.06 (d, *J* = 1.0 Hz, 3H), 1.25 (t, *J* = 7.1 Hz, 1H). ^{13}C NMR (75 MHz, $CDCl_3$) δ 168.2, 166.3, 150.3, 149.2, 136.5, 135.6, 134.4, 133.4, 132.3,

129.5, 129.2, 128.3, 127.7, 124.4, 116.5, 114.0, 106.4, 13.1, 11.5. HRMS (ESI, CH_3OH) $C_{21}H_{17}N_3O_2S$ Calc.: 375.1041 Found: 398.0949 $[M+Na]^+$.

1.9 References

1. http://seer.cancer.gov.

2. Lepourcelet, M.; Chen, Y. N.; France, D. S.; Wang, H.; Crews, P.; Petersen, F.; Bruseo, C.; Wood, A. W.; Shivdasani, R. A., Small-molecule antagonists of the oncogenic Tcf/beta-catenin protein complex. *Cancer Cell* **2004,** *5* (1), 91-102.

3. Nusse, R.; Varmus, H., Three decades of Wnts: a personal perspective on how a scientific field developed. *EMBO J* **2012,** *31* (12), 2670-84.

4. Korteweg, R., On the manner in which the disposition to carcinoma of the mammary gland is inherited in mice. *Genetica* **1936,** *18* (3), 350-371.

5. Bittner, J. J., Some possible effects of nursing on the mammary gland tumor incidence in mice. *Science (New York, N.Y.)* **1936,** *84* (2172), 162-162.

6. Lyons, M. J.; Moore, D. H., Purification of the mouse mammary tumour virus. *Nature* **1962,** *194*, 1141-2.

7. Bishop, J. M., Cellular oncogenes and retroviruses. *Annual Review of Biochemistry* **1983,** *52*, 301-54.

8. Sharma, R. P.; Chopra, V. L., Effect of the wingless (wg1) mutation on wing and haltere development in Drosophila melanogaster. *Developmental Biology* **1976,** *48* (2), 461-5.

9. Nusslein-Volhard, C.; Wieschaus, E., Mutations affecting segment number and polarity in Drosophila. *Nature* **1980,** *287* (5785), 795-801.

10. Nusse, R.; Varmus, H. E., Many tumors induced by the mouse mammary tumor virus contain a provirus integrated in the same region of the host genome. *Cell* **1982,** *31* (1), 99-109.

11. van Ooyen, A.; Nusse, R., Structure and nucleotide sequence of the putative mammary oncogene int-1; proviral insertions leave the protein-encoding domain intact. *Cell* **1984,** *39* (1), 233-40.

12. Dickson, C.; Smith, R.; Brookes, S.; Peters, G., Tumorigenesis by mouse mammary tumor virus: proviral activation of a cellular gene in the common integration region

int-2. *Cell* **1984,** *37* (2), 529-36.

13. Bishop, J. M.; Varmus, H., Functions and origins of retroviral transforming genes. In *RNA Tumor Viruses*, Weis, R.; Teich, N.; Varmus, H.; Coffin, J. M., Eds. Cold Spring Harbor: Cold Spring Harbor Laboratory: 1985; pp 249-356.

14. Fung, Y. K.; Shackleford, G. M.; Brown, A. M.; Sanders, G. S.; Varmus, H. E., Nucleotide sequence and expression in vitro of cDNA derived from mRNA of int-1, a provirally activated mouse mammary oncogene. *Molecular and Cellular Biology* **1985,** *5* (12), 3337-44.

15. van Ooyen, A.; Kwee, V.; Nusse, R., The nucleotide sequence of the human int-1 mammary oncogene; evolutionary conservation of coding and non-coding sequences. *EMBO J* **1985,** *4* (11), 2905-9.

16. Brown, A. M.; Wildin, R. S.; Prendergast, T. J.; Varmus, H. E., A retrovirus vector expressing the putative mammary oncogene int-1 causes partial transformation of a mammary epithelial cell line. *Cell* **1986,** *46* (7), 1001-9.

17. Rijsewijk, F.; Schuermann, M.; Wagenaar, E.; Parren, P.; Weigel, D.; Nusse, R., The Drosophila homolog of the mouse mammary oncogene int-1 is identical to the segment polarity gene wingless. *Cell* **1987,** *50* (4), 649-57.

18. Baker, N. E., Molecular cloning of sequences from wingless, a segment polarity gene in Drosophila: the spatial distribution of a transcript in embryos. *EMBO J* **1987,** *6* (6), 1765-73.

19. Gallahan, D.; Callahan, R., Mammary tumorigenesis in feral mice: identification of a new int locus in mouse mammary tumor virus (Czech II)-induced mammary tumors. *Journal of Virology* **1987,** *61* (1), 66-74.

20. Tsukamoto, A. S.; Grosschedl, R.; Guzman, R. C.; Parslow, T.; Varmus, H. E., Expression of the int-1 gene in transgenic mice is associated with mammary gland hyperplasia and adenocarcinomas in male and female mice. *Cell* **1988,** *55* (4), 619-25.

21. Ozawa, M.; Baribault, H.; Kemler, R., The cytoplasmic domain of the cell adhesion molecule uvomorulin associates with three independent proteins structurally related in different species. *EMBO J* **1989,** *8* (6), 1711-7.

22. McMahon, A. P.; Moon, R. T., Ectopic expression of the proto-oncogene int-1 in Xenopus embryos leads to duplication of the embryonic axis. *Cell* **1989,** *58* (6), 1075-84.

23. Roelink, H.; Wagenaar, E.; Lopes da Silva, S.; Nusse, R., Wnt-3, a gene activated by proviral insertion in mouse mammary tumors, is homologous to int-1/Wnt-1 and

is normally expressed in mouse embryos and adult brain. *Proceedings of the National Academy of Sciences of the United States of America* **1990,** *87* (12), 4519-23.

24. Gavin, B. J.; McMahon, J. A.; McMahon, A. P., Expression of multiple novel Wnt-1/int-1-related genes during fetal and adult mouse development. *Genes & Development* **1990,** *4* (12b), 2319-32.

25. Takeichi, M., Cadherins: a molecular family important in selective cell-cell adhesion. *Annual Review of Biochemistry* **1990,** *59*, 237-52.

26. Groden, J.; Thliveris, A.; Samowitz, W.; Carlson, M.; Gelbert, L.; Albertsen, H.; Joslyn, G.; Stevens, J.; Spirio, L.; Robertson, M.; et al., Identification and characterization of the familial adenomatous polyposis coli gene. *Cell* **1991,** *66* (3), 589-600.

27. Kinzler, K. W.; Nilbert, M. C.; Su, L. K.; Vogelstein, B.; Bryan, T. M.; Levy, D. B.; Smith, K. J.; Preisinger, A. C.; Hedge, P.; McKechnie, D.; et al., Identification of FAP locus genes from chromosome 5q21. *Science (New York, N.Y.)* **1991,** *253* (5020), 661-5.

28. Peifer, M.; Wieschaus, E., The segment polarity gene armadillo encodes a functionally modular protein that is the Drosophila homolog of human plakoglobin. *Cell* **1990,** *63* (6), 1167-76.

29. Nusse, R.; Brown, A.; Papkoff, J.; Scambler, P.; Shackleford, G.; McMahon, A.; Moon, R.; Varmus, H., A new nomenclature for int-1 and related genes: the Wnt gene family. *Cell* **1991,** *64* (2), 231.

30. Nusse, R.; Varmus, H. E., Wnt genes. *Cell* **1992,** *69* (7), 1073-87.

31. Siegfried, E.; Chou, T. B.; Perrimon, N., wingless signaling acts through zeste-white 3, the Drosophila homolog of glycogen synthase kinase-3, to regulate engrailed and establish cell fate. *Cell* **1992,** *71* (7), 1167-79.

32. Rubinfeld, B.; Souza, B.; Albert, I.; Muller, O.; Chamberlain, S. H.; Masiarz, F. R.; Munemitsu, S.; Polakis, P., Association of the APC gene product with beta-catenin. *Science (New York, N.Y.)* **1993,** *262* (5140), 1731-4.

33. Peifer, M.; Orsulic, S.; Pai, L. M.; Loureiro, J., A model system for cell adhesion and signal transduction in Drosophila. *Development (Cambridge, England). Supplement* **1993**, 163-76.

34. Peifer, M.; Pai, L. M.; Casey, M., Phosphorylation of the Drosophila adherens junction protein Armadillo: roles for wingless signal and zeste-white 3 kinase. *Developmental Biology* **1994,** *166* (2), 543-56.

35. van Leeuwen, F.; Samos, C. H.; Nusse, R., Biological activity of soluble wingless protein in cultured Drosophila imaginal disc cells. *Nature* **1994,** *368* (6469), 342-4.

36. He, T. C.; Sparks, A. B.; Rago, C.; Hermeking, H.; Zawel, L.; da Costa, L. T.; Morin, P. J.; Vogelstein, B.; Kinzler, K. W., Identification of c-MYC as a target of the APC pathway. *Science (New York, N.Y.)* **1998,** *281* (5382), 1509-12.

37. Funayama, N.; Fagotto, F.; McCrea, P.; Gumbiner, B. M., Embryonic axis induction by the armadillo repeat domain of beta-catenin: evidence for intracellular signaling. *The Journal of Cell Biology* **1995,** *128* (5), 959-68.

38. Rubinfeld, B.; Albert, I.; Porfiri, E.; Fiol, C.; Munemitsu, S.; Polakis, P., Binding of GSK3beta to the APC-beta-catenin complex and regulation of complex assembly. *Science (New York, N.Y.)* **1996,** *272* (5264), 1023-6.

39. Molenaar, M.; van de Wetering, M.; Oosterwegel, M.; Peterson-Maduro, J.; Godsave, S.; Korinek, V.; Roose, J.; Destree, O.; Clevers, H., XTcf-3 transcription factor mediates beta-catenin-induced axis formation in Xenopus embryos. *Cell* **1996,** *86* (3), 391-9.

40. Bhanot, P.; Brink, M.; Samos, C. H.; Hsieh, J. C.; Wang, Y.; Macke, J. P.; Andrew, D.; Nathans, J.; Nusse, R., A new member of the frizzled family from Drosophila functions as a Wingless receptor. *Nature* **1996,** *382* (6588), 225-30.

41. Finch, P. W.; He, X.; Kelley, M. J.; Uren, A.; Schaudies, R. P.; Popescu, N. C.; Rudikoff, S.; Aaronson, S. A.; Varmus, H. E.; Rubin, J. S., Purification and molecular cloning of a secreted, Frizzled-related antagonist of Wnt action. *Proceedings of the National Academy of Sciences of the United States of America* **1997,** *94* (13), 6770-5.

42. Leyns, L.; Bouwmeester, T.; Kim, S. H.; Piccolo, S.; De Robertis, E. M., Frzb-1 is a secreted antagonist of Wnt signaling expressed in the Spemann organizer. *Cell* **1997,** *88* (6), 747-56.

43. Rattner, A.; Hsieh, J. C.; Smallwood, P. M.; Gilbert, D. J.; Copeland, N. G.; Jenkins, N. A.; Nathans, J., A family of secreted proteins contains homology to the cysteine-rich ligand-binding domain of frizzled receptors. *Proceedings of the National Academy of Sciences of the United States of America* **1997,** *94* (7), 2859-63.

44. Riese, J.; Yu, X.; Munnerlyn, A.; Eresh, S.; Hsu, S. C.; Grosschedl, R.; Bienz, M., LEF-1, a nuclear factor coordinating signaling inputs from wingless and decapentaplegic. *Cell* **1997,** *88* (6), 777-87.

45. Behrens, J.; Jerchow, B. A.; Wurtele, M.; Grimm, J.; Asbrand, C.; Wirtz, R.; Kuhl, M.; Wedlich, D.; Birchmeier, W., Functional interaction of an axin homolog, conductin, with beta-catenin, APC, and GSK3beta. *Science (New York, N.Y.)* **1998,**

280 (5363), 596-9.

46. Ikeda, S.; Kishida, S.; Yamamoto, H.; Murai, H.; Koyama, S.; Kikuchi, A., Axin, a negative regulator of the Wnt signaling pathway, forms a complex with GSK-3beta and beta-catenin and promotes GSK-3beta-dependent phosphorylation of beta-catenin. *EMBO J* **1998,** *17* (5), 1371-84.

47. Pinson, K. I.; Brennan, J.; Monkley, S.; Avery, B. J.; Skarnes, W. C., An LDL-receptor-related protein mediates Wnt signalling in mice. *Nature* **2000,** *407* (6803), 535-8.

48. Tamai, K.; Zeng, X.; Liu, C.; Zhang, X.; Harada, Y.; Chang, Z.; He, X., A mechanism for Wnt coreceptor activation. *Molecular Cell* **2004,** *13* (1), 149-56.

49. Mao, B.; Wu, W.; Li, Y.; Hoppe, D.; Stannek, P.; Glinka, A.; Niehrs, C., LDL-receptor-related protein 6 is a receptor for Dickkopf proteins. *Nature* **2001,** *411* (6835), 321-5.

50. Clevers, H.; Nusse, R., Wnt/b-catenin signaling and disease. *Cell* **2012,** *149* (6), 1192-1205.

51. Nusse, R.; Ooyen, A. V.; Cox, D.; Fung, Y. K. T.; Varmus, H., Mode of proviral activation of a putative mammary oncogene (int-1) on mouse chromosome 15. *Nature* **1984,** *307*, 131-136.

52. de La Coste, A.; Romagnolo, B.; Billuart, P.; Renard, C. A.; Buendia, M. A.; Soubrane, O.; Fabre, M.; Chelly, J.; Beldjord, C.; Kahn, A.; Perret, C., Somatic mutations of the beta-catenin gene are frequent in mouse and human hepatocellular carcinomas. *Proceedings of the National Academy of Sciences of the United States of America* **1998,** *95* (15), 8847-51.

53. Korinek, V.; Barker, N.; Morin, P. J.; van Wichen, D.; de Weger, R.; Kinzler, K. W.; Vogelstein, B.; Clevers, H., Constitutive transcriptional activation by a beta-catenin-Tcf complex in APC-/- colon carcinoma. *Science (New York, N.Y.)* **1997,** *275* (5307), 1784-7.

54. Morin, P. J.; Sparks, A. B.; Korinek, V.; Barker, N.; Clevers, H.; Vogelstein, B.; Kinzler, K. W., Activation of beta-catenin-Tcf signaling in colon cancer by mutations in beta-catenin or APC. *Science (New York, N.Y.)* **1997,** *275* (5307), 1787-90.

55. Rubinfeld, B.; Robbins, P.; El-Gamil, M.; Albert, I.; Porfiri, E.; Polakis, P., Stabilization of beta-catenin by genetic defects in melanoma cell lines. *Science (New York, N.Y.)* **1997,** *275* (5307), 1790-2.

56. Satoh, S.; Daigo, Y.; Furukawa, Y.; Kato, T.; Miwa, N.; Nishiwaki, T.; Kawasoe,

T.; Ishiguro, H.; Fujita, M.; Tokino, T.; Sasaki, Y.; Imaoka, S.; Murata, M.; Shimano, T.; Yamaoka, Y.; Nakamura, Y., AXIN1 mutations in hepatocellular carcinomas, and growth suppression in cancer cells by virus-mediated transfer of AXIN1. *Nature Genetics* **2000,** *24* (3), 245-50.

57. Barker, N.; Clevers, H., Mining the Wnt pathway for cancer therapeutics. *Nature Reviews. Drug Discovery* **2006,** *5* (12), 997-1014.

58. Takahashi-Yanaga, F.; Kahn, M., Targeting Wnt signaling: can we safely eradicate cancer stem cells? *Clinical Cancer Research : an Official Journal of the American Association for Cancer Research* **2010,** *16* (12), 3153-62.

59. van Noort, M.; Clevers, H., TCF transcription factors, mediators of Wnt-signaling in development and cancer. *Developmental Biology* **2002,** *244* (1), 1-8.

60. Shtutman, M.; Zhurinsky, J.; Simcha, I.; Albanese, C.; D'Amico, M.; Pestell, R.; Ben-Ze'ev, A., The cyclin D1 gene is a target of the beta-catenin/LEF-1 pathway. *Proceedings of the National Academy of Sciences of the United States of America* **1999,** *96* (10), 5522-7.

61. Tetsu, O.; McCormick, F., Beta-catenin regulates expression of cyclin D1 in colon carcinoma cells. *Nature* **1999,** *398* (6726), 422-6.

62. van de Wetering, M.; Sancho, E.; Verweij, C.; de Lau, W.; Oving, I.; Hurlstone, A.; van der Horn, K.; Batlle, E.; Coudreuse, D.; Haramis, A. P.; Tjon-Pon-Fong, M.; Moerer, P.; van den Born, M.; Soete, G.; Pals, S.; Eilers, M.; Medema, R.; Clevers, H., The beta-catenin/TCF-4 complex imposes a crypt progenitor phenotype on colorectal cancer cells. *Cell* **2002,** *111* (2), 241-50.

63. He, B.; Reguart, N.; You, L.; Mazieres, J.; Xu, Z.; Lee, A. Y.; Mikami, I.; McCormick, F.; Jablons, D. M., Blockade of Wnt-1 signaling induces apoptosis in human colorectal cancer cells containing downstream mutations. *Oncogene* **2005,** *24* (18), 3054-8.

64. Polakis, P., Drugging Wnt signalling in cancer. *EMBO J* **2012,** *31* (12), 2737-46.

65. Anastas, J. N.; Moon, R. T., WNT signalling pathways as therapeutic targets in cancer. *Nature Reviews. Cancer* **2013,** *13* (1), 11-26.

66. Kim, J. S.; Crooks, H.; Foxworth, A.; Waldman, T., Proof-of-principle: oncogenic beta-catenin is a valid molecular target for the development of pharmacological inhibitors. *Molecular Cancer Therapeutics* **2002,** *1* (14), 1355-9.

67. Ashihara, E.; Kawata, E.; Nakagawa, Y.; Shimazaski, C.; Kuroda, J.; Taniguchi, K.; Uchiyama, H.; Tanaka, R.; Yokota, A.; Takeuchi, M.; Kamitsuji, Y.; Inaba, T.; Taniwaki, M.; Kimura, S.; Maekawa, T., beta-catenin small interfering RNA

successfully suppressed progression of multiple myeloma in a mouse model. *Clinical Cancer Research : an Official Journal of the American Association for Cancer Research* **2009,** *15* (8), 2731-8.

68. Scholer-Dahirel, A.; Schlabach, M. R.; Loo, A.; Bagdasarian, L.; Meyer, R.; Guo, R.; Woolfenden, S.; Yu, K. K.; Markovits, J.; Killary, K.; Sonkin, D.; Yao, Y. M.; Warmuth, M.; Sellers, W. R.; Schlegel, R.; Stegmeier, F.; Mosher, R. E.; McLaughlin, M. E., Maintenance of adenomatous polyposis coli (APC)-mutant colorectal cancer is dependent on Wnt/beta-catenin signaling. *Proceedings of the National Academy of Sciences of the United States of America* **2011,** *108* (41), 17135-40.

69. Yeung, J.; Esposito, M. T.; Gandillet, A.; Zeisig, B. B.; Griessinger, E.; Bonnet, D.; So, C. W., Beta-Catenin mediates the establishment and drug resistance of MLL leukemic stem cells. *Cancer Cell* **2010,** *18* (6), 606-18.

70. Weber, B. N.; Chi, A. W.; Chavez, A.; Yashiro-Ohtani, Y.; Yang, Q.; Shestova, O.; Bhandoola, A., A critical role for TCF-1 in T-lineage specification and differentiation. *Nature* **2011,** *476* (7358), 63-8.

71. Wells, J. A.; McClendon, C. L., Reaching for high-hanging fruit in drug discovery at protein-protein interfaces. *Nature* **2007,** *450* (7172), 1001-9.

72. Kuriyan, J.; Eisenberg, D., The origin of protein interactions and allostery in colocalization. *Nature* **2007,** *450* (7172), 983-90.

73. Sun, J.; Weis, W. I., Biochemical and structural characterization of beta-catenin interactions with nonphosphorylated and CK2-phosphorylated Lef-1. *Journal of Molecular Biology* **2011,** *405* (2), 519-30.

74. Knapp, S.; Zamai, M.; Volpi, D.; Nardese, V.; Avanzi, N.; Breton, J.; Plyte, S.; Flocco, M.; Marconi, M.; Isacchi, A.; Caiolfa, V. R., Thermodynamics of the high-affinity interaction of TCF4 with beta-catenin. *Journal of Molecular Biology* **2001,** *306* (5), 1179-89.

75. Choi, H. J.; Huber, A. H.; Weis, W. I., Thermodynamics of beta-catenin-ligand interactions: the roles of the N- and C-terminal tails in modulating binding affinity. *The Journal of Biological Chemistry* **2006,** *281* (2), 1027-38.

76. Ha, N. C.; Tonozuka, T.; Stamos, J. L.; Choi, H. J.; Weis, W. I., Mechanism of phosphorylation-dependent binding of APC to beta-catenin and its role in beta-catenin degradation. *Molecular Cell* **2004,** *15* (4), 511-21.

77. Choi, H. J.; Gross, J. C.; Pokutta, S.; Weis, W. I., Interactions of plakoglobin and beta-catenin with desmosomal cadherins: basis of selective exclusion of alpha- and beta-catenin from desmosomes. *The Journal of Biological Chemistry* **2009,** *284* (46),

31776-88.

78. Sampietro, J.; Dahlberg, C. L.; Cho, U. S.; Hinds, T. R.; Kimelman, D.; Xu, W., Crystal structure of a beta-catenin/BCL9/Tcf4 complex. *Molecular Cell* **2006,** *24* (2), 293-300.

79. Trosset, J. Y.; Dalvit, C.; Knapp, S.; Fasolini, M.; Veronesi, M.; Mantegani, S.; Gianellini, L. M.; Catana, C.; Sundstrom, M.; Stouten, P. F.; Moll, J. K., Inhibition of protein-protein interactions: the discovery of druglike beta-catenin inhibitors by combining virtual and biophysical screening. *Proteins* **2006,** *64* (1), 60-7.

80. Tian, W.; Han, X.; Yan, M.; Xu, Y.; Duggineni, S.; Lin, N.; Luo, G.; Li, Y. M.; Han, X.; Huang, Z.; An, J., Structure-based discovery of a novel inhibitor targeting the beta-catenin/Tcf4 interaction. *Biochemistry* **2012,** *51* (2), 724-31.

81. Gonsalves, F. C.; Klein, K.; Carson, B. B.; Katz, S.; Ekas, L. A.; Evans, S.; Nagourney, R.; Cardozo, T.; Brown, A. M.; DasGupta, R., An RNAi-based chemical genetic screen identifies three small-molecule inhibitors of the Wnt/wingless signaling pathway. *Proceedings of the National Academy of Sciences of the United States of America* **2011,** *108* (15), 5954-63.

82. von Kries, J. P.; Winbeck, G.; Asbrand, C.; Schwarz-Romond, T.; Sochnikova, N.; Dell'Oro, A.; Behrens, J.; Birchmeier, W., Hot spots in beta-catenin for interactions with LEF-1, conductin and APC. *Nature Structural Biology* **2000,** *7* (9), 800-7.

83. Baell, J. B.; Holloway, G. A., New substructure filters for removal of pan assay interference compounds (PAINS) from screening libraries and for their exclusion in bioassays. *Journal of Medicinal Chemistry* **2010,** *53* (7), 2719-40.

84. Zhang, M.; Catrow, J. L.; Ji, H., High-Throughput Selectivity Assays for Small-Molecule Inhibitors of beta-Catenin/T-Cell Factor Protein-Protein Interactions. *ACS Medicinal Chemistry Letters* **2013,** *4* (2), 306-11.

85. Grossmann, T. N.; Yeh, J. T.; Bowman, B. R.; Chu, Q.; Moellering, R. E.; Verdine, G. L., Inhibition of oncogenic Wnt signaling through direct targeting of beta-catenin. *Proceedings of the National Academy of Sciences of the United States of America* **2012,** *109* (44), 17942-7.

86. Zhang, M.; Huang, Z.; Yu, B.; Ji, H., New homogeneous high-throughput assays for inhibitors of beta-catenin/Tcf protein-protein interactions. *Analytical Biochemistry* **2012,** *424* (1), 57-63.

87. Hasegawa, K.; Yamada, K.; Sasase, R.; Miyazaki, R.; Kikuchi, A.; Yagi, M., Direct measurements of absolute concentration and lifetime of singlet oxygen in the gas phase by electron paramagnetic resonance. *Chemical Physics Letters* **2008,** *457* (4–6), 312-314.

88. Yu, B.; Huang, Z.; Zhang, M.; Dillard, D. R.; Ji, H., Rational design of small-molecule inhibitors for beta-catenin/T-cell factor protein-protein interactions by bioisostere replacement. *ACS Chemical Biology* **2013,** *8* (3), 524-9.

89. Dakin, H. D.; West, R., A general reaction of amino acids. *The Journal of Biological Chemistry* **1928,** *78*, 91-104.

90. Moon, H.-S.; Yoo, M.-H.; Kim, S.-H.; Lim, J.-I.; Son, M.-H.; Kim, M.-K.; Shin, C.-Y.; Kim, J.-K.; Park, S.-K.; Chae, Y.-N.; Shim, H.-J.; Jeon, S.-H.; Kim, H.-S.; Wie, G.-T.; Kim, D.-H.; Lee, B.-K.; Park, C.-S.; Ahn, B.-N.; Kim, E.; Bae, M.-H.; Shin, Y.-A.; Hur, Y.; Lee, C.-H.; Choi, H.-H.; Kim, B.; Chong, W. Novel phenylpropionic acid derivatives as peroxisome proliferator-activated gamma receptor modulators, method of preparing the same, and pharmaceutical composition comprising the same. WO2008108602A1, 2008.

91. Yamane, T.; Mitsudera, H.; Shundoh, T., A new approach to the synthesis of 2-aryl-4-halomethyl-5-methyl-1,3-oxazoles by highly regioselective direct halogenation with NBS or NSC/MeCN. *Synthesis* **2004,** (17), 2825-2832.

92. Knoevengal, E., Condensation von malonsäure mit aromatiachen aldehyden durch ammoniak und amine. *Berichte der Deutschen Chemischen Gesellschaft* **1898,** *31* (3), 2596-2619.

93. Sarkis, M.; Tran, D. N.; Kolb, S.; Miteva, M. A.; Villoutreix, B. O.; Garbay, C.; Braud, E., Design and synthesis of novel bis-thiazolone derivatives as micromolar CDC25 phosphatase inhibitors: effect of dimerisation on phosphatase inhibition. *Bioorg. Med. Chem. Lett.* **2012,** *22* (24), 7345-7350.

94. Macdonald, J. E.; Hysell, M. K.; Yu, D.; Li, H.; Wong-Staal, F. Preparation of quinolinium salts as anticancer drugs. WO2006078754A1, 2006.

95. Sahu, M.; Garnaik, B. K.; Behera, R., Influence of substituents on the synthesis of thiazolidinones. *Indian J. Chem., Sect. B* **1987,** *26B* (8), 779-81.

96. Villoutreix, B. O.; Kuenemann, M. A.; Poyet, J.-L.; Bruzzoni-Giovanelli, H.; Labbé, C.; Lagorce, D.; Sperandio, O.; Miteva, M. A., Drug-like protein–protein interaction modulators: challenges and opportunities for drug discovery and chemical biology. *Molecular Informatics* **2014,** *33* (6-7), 414-437.

97. Clackson, T.; Wells, J. A., A hot spot of binding energy in a hormone-receptor interface. *Science (New York, N.Y.)* **1995,** *267* (5196), 383-6.

98. Graham, T. A.; Weaver, C.; Mao, F.; Kimelman, D.; Xu, W., Crystal structure of a beta-catenin/Tcf complex. *Cell* **2000,** *103* (6), 885-96.

99. Graham, T. A.; Ferkey, D. M.; Mao, F.; Kimelman, D.; Xu, W., Tcf4 can

specifically recognize beta-catenin using alternative conformations. *Nature Structural Biology* **2001,** *8* (12), 1048-52.

100. Gail, R.; Frank, R.; Wittinghofer, A., Systematic peptide array-based delineation of the differential beta-catenin interaction with Tcf4, E-cadherin, and adenomatous polyposis coli. *The Journal of Biological Chemistry* **2005,** *280* (8), 7107-17.

101. Poy, F.; Lepourcelet, M.; Shivdasani, R. A.; Eck, M. J., Structure of a human Tcf4-beta-catenin complex. *Nature Structural Biology* **2001,** *8* (12), 1053-7.

102. Omer, C. A.; Miller, P. J.; Diehl, R. E.; Kral, A. M., Identification of Tcf4 residues involved in high-affinity beta-catenin binding. *Biochemical and Biophysical Research Communications* **1999,** *256* (3), 584-90.

103. Fasolini, M.; Wu, X.; Flocco, M.; Trosset, J. Y.; Oppermann, U.; Knapp, S., Hot spots in Tcf4 for the interaction with beta-catenin. *The Journal of Biological Chemistry* **2003,** *278* (23), 21092-8.

104. Buckley, D. L.; Van Molle, I.; Gareiss, P. C.; Tae, H. S.; Michel, J.; Noblin, D. J.; Jorgensen, W. L.; Ciulli, A.; Crews, C. M., Targeting the von Hippel-Lindau E3 ubiquitin ligase using small molecules to disrupt the VHL/HIF-1alpha interaction. *Journal of the American Chemical Society* **2012,** *134* (10), 4465-8.

105. Whitby, L. R.; Boger, D. L., Comprehensive peptidomimetic libraries targeting protein-protein interactions. *Accounts of Chemical Research* **2012,** *45* (10), 1698-709.

CHAPTER 2

TARGETING THE TCF4 GANDE BINDING SITE TO SELECTIVELY DISRUPT *β*–CATENIN/T-CELL FACTOR PROTEIN-PROTEIN INTERACTIONS USING A PEPTIDOMIMETIC STRATEGY

2.1 Introduction

In the postgenomic era, a myriad of functional proteins have been characterized. Inside of the cell, there is a complex network of protein-protein interactions (PPIs). The protein tends to use the same interface to interact with the various other proteins in different organelles or cellular environments. Under pathological conditions, PPIs lose the compartmental constraint, and unexpected PPIs take place, often leading to disease. One challenge of biomedical research is to design a small molecule to modulate a specific protein-protein interface while leaving the other related PPIs unaffected. The selective disruption of *β*-catenin/Tcf PPIs, a key downstream effector of the canonical Wnt signaling pathway, represents such a case.

The canonical Wnt signaling pathway plays a fundamental role in directing cell proliferation, differentiation, and survival.[1] *β* -Catenin is the key mediator of this signaling pathway. The mutations in the regulatory genes of the canonical Wnt signaling pathway, such as those that encode Adenomatous Polyposis Coli (APC), Axin, and/or the *N*-terminal phosphorylation sites of *β*-catenin, result in the stabilization of *β*-catenin and an increased

level of nuclear *β*-catenin. Nuclear *β*-catenin forms a complex with T-cell factor (Tcf)/lymphoid enhancer factor (Lef) to induce the overexpression of cell proliferation, migration, and survival genes, such as *cyclin D1*, c-myc, and survivin. Canonical Wnt signaling can also be activated through epigenetic silencing of Wnt antagonist genes or the autocrine activation of Wnt ligands, frizzled (Fz), and disheveled (Dvl). The hyperactivation of the canonical Wnt signaling pathway is strongly implicated in the initiation and progression of many cancers and fibrosis. In addition, canonical Wnt signaling is aberrantly over-activated in cancer stem cells, which drive tumor growth, seed metastases, and cause cancer recurrence after remission.[2-4] Significant efforts have been made to identify inhibitors of the canonical Wnt signaling pathway.[5-7] Inhibitors of the upstream sites are less desirable because they can cause cross-regulatory effects on the *β*-catenin independent Wnt pathways or perturb the function of *β*-catenin in cell-cell adhesion. Thus, it would be better to identify inhibitors for the downstream sites of the canonical Wnt signaling pathway. The formation of the *β*-catenin/Tcf complex in the cell nucleus is the penultimate step of canonical Wnt signaling. Transcriptional over-activation of canonical Wnt target genes is solely dependent on the formation of this complex. Therefore, selective inhibition of *β*-catenin/Tcf PPIs represents an appealing therapeutic target. It is important that the inhibitor binds to coactivator *β*-catenin rather than transcriptional factor Tcf, because Tcf is essential for the other signaling pathways,[8] while *β*-catenin is specific for canonical Wnt signaling.

The crystal structures of *β*-catenin in complexes with Xenopus Tcf3,[9] human Tcf4,[10-12] mouse Lef1,[13] E-cadherin,[14] and APC[15-16] have been solved. A comparison of these structures reveals that *β*-catenin uses the same armadillo repeats to bind Tcf (human

Tcf4 residues 7−51), cadherin (human E-cadherin residues 819−873), and APC (residues 1477−1519 of human APC 20-amino acid repeat 3, APC-R3). The binding modes of Tcf, cadherin, and APC in this region are identical. Biochemical analyses confirm that the binding modes of Tcf, cadherin, and APC to *β*-catenin are mutually exclusive.[17-20] Selectivity becomes a major challenge when designing inhibitors for *β*-catenin/Tcf interactions. The inhibitor cannot block the *β*-catenin/cadherin interaction because this interaction is essential for the integrity of epithelial junctions. The blockage of this protein-protein complex alters cell-cell adhesion and impairs the functions of normal stem cells. The inhibitor cannot block the *β*-catenin/APC interaction either because this interaction is critical for *β*-catenin degradation. The blockage of this protein-protein complex leads to a larger pool of free *β*-catenin, and more *β*-catenin enters the cell nuclei to over-activate the transcription of Wnt target genes, thereby promoting new cancer formation. In addition, the dissociation constant (K_d) of the *β*-catenin/Tcf complex is 7−10 nM[13, 21] and much lower than those of the *β*-catenin/cadherin (K_d: 41−82 nM)[20] and *β*-catenin/APC (K_d: 0.6−3.1 μM for APC-R3, the APC repeat with the highest binding affinity) complexes,[15, 21] which further complicates the design of selective *β*-catenin/Tcf inhibitors. The disruption of *β*-catenin/Tcf interactions has been thought undruggable for these reasons.[5, 12]

Although extensive high-throughput screening (HTS) has been performed to identify small-molecule inhibitors for *β*-catenin/Tcf PPIs, only four examples have been reported (**Figure 2.1**).[22-25] However, the exact binding modes of these HTS hits remain unknown and hence have prohibited their optimization. The selectivity of these inhibitors for *β*-catenin/Tcf over *β*-catenin/E-cadherin and *β*-catenin/APC interactions was low.[26] Hydrocarbon-stapled Axin α-helix, **aStAx-35**, was identified to bind to the Axin-binding

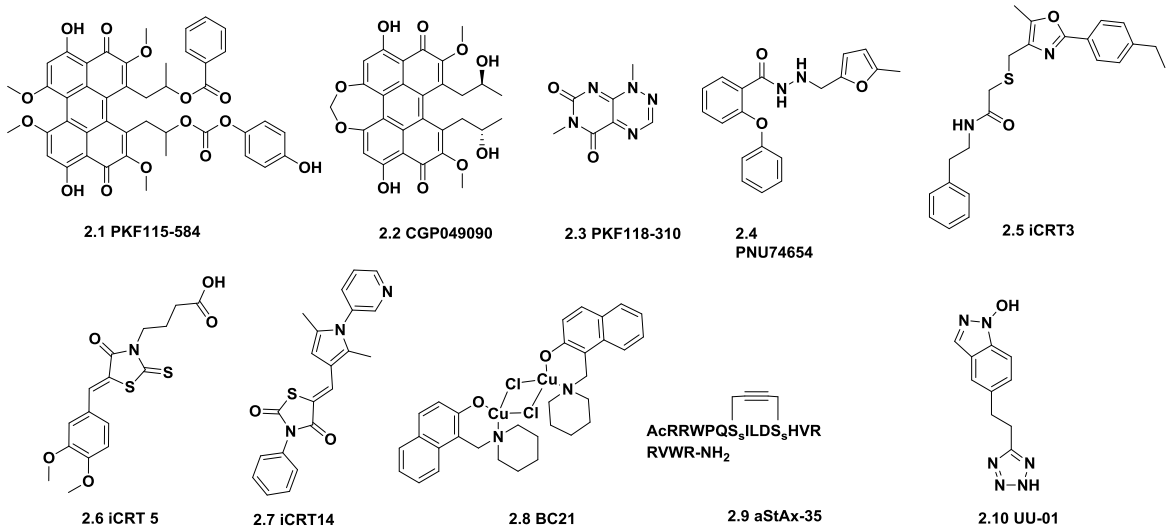


Figure 2.1 Chemical structures of known *β*-catenin/Tcf inhibitors.

site of *β*-catenin and inhibit *β*-catenin/Tcf interactions.[27] Axin is the scaffolding protein for *β*-catenin phosphorylation. *β*-Catenin/Axin interactions maintain a low level of free *β*-catenin in the cytoplasm. The use of these hydrocarbon-stapled peptides has the potential to elevate the concentrations of *β*-catenin in normal cells. Yu *et al*. reported a rational design of small molecule inhibitor **UU-01** by bioisostere replacement.[28]

In this study, we demonstrated that targeting the Tcf4 G^{13}ANDE^{17} binding site can selectively inhibit *β*-catenin/Tcf PPIs. Small-molecule inhibitors with excellent cell-based activity and high selectivity for *β*-catenin/Tcf over *β*-catenin/cadherin were identified.

2.2 Results and Discussion

2.2.1 Identification of a Selective Binding Site for β-Catenin/Tcf over β-Catenin/E-Cadherin and β-Catenin/APC Interactions

A careful inspection of the crystal structures of β-catenin reveals that the binding modes of the sequences flanking Tcf4 D16, E-cadherin D830, and APC-R3 D1486 have subtle but critical differences. A kink between armadillo repeats 9 and 10 of β-catenin forms a binding site (**Figure 2.2**, hot spot 1[26]). Tcf4 G[13]ANDE[17] binds to this site. There are four well-defined pockets for the Tcf4 G[13]ANDE[17] binding site: pocket A is lined with C429, N430, K435, H470, S473, and R474; pocket B is deep and lined with V511, C573, L539, I569, I507, and the side chain carbons of K508; pocket C is a surface pocket defined by L519 and I579; and pocket D is another surface pocket that is lined with C466, E462, L506, A509, and the side chain carbons of K508. Pockets A and B are connected to pocket C. Pockets B and D are connected by K508 and R469.

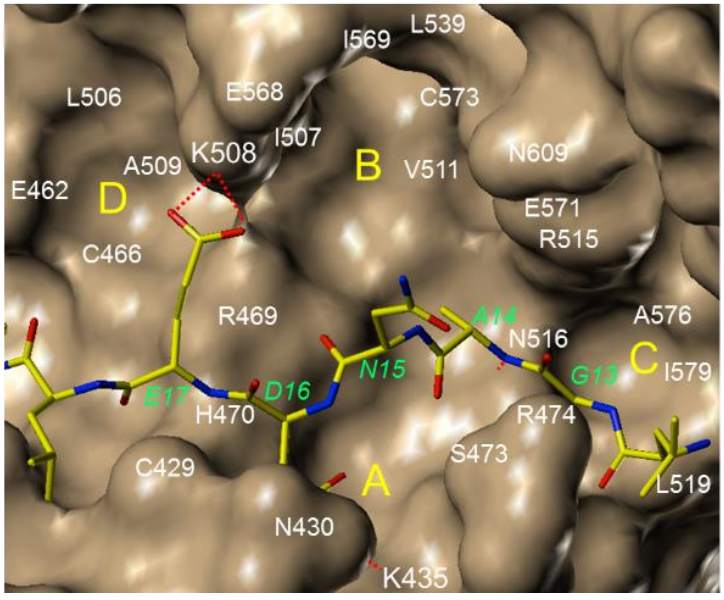


Figure 2.2. Pockets A–D of the Tcf GANDE binding site. Tcf4 G[13]ANDE[17] is shown yellow as the reference.

Residue K435 in pocket A forms a salt bridge with D16 of human Tcf4, D830 of human E-cadherin, and D1486 of human APC-R3. This interaction is essential for anchoring Tcf/Lef,[9, 11, 18, 21, 29-31] cadherin,[9, 30] and APC[13, 32] in place. However, the binding modes adjacent to this salt bridge are different among the three protein–protein complexes. Tcf4 E17 forms a salt bridge with *β*-catenin K508. *β*-Catenin K508A or Tcf4 E17A mutation significantly decreases the *β*-catenin/Tcf affinity.[18, 21, 28, 30] Our previous study demonstrated that Tcf4 $D^{16}E^{17}$ alone could effectively disrupt *β*-catenin/Tcf interactions with a K_i value of 396.59 μM.[28] The mimicking of Tcf4 $D^{16}E^{17}$ by bioisostere replacement generated a potent *β*-catenin/Tcf inhibitor, **UU-01**, with a K_i value of 3.14 μM.

On the other hand, the residues in E-cadherin and APC-R3 that correspond to Tcf4 E17 are S831 and T1487, respectively. These two residues do not form a direct bonding with *β*-catenin. The backbone amide of Tcf4 A14 forms an H-bond with the side chain carboxamide of N516. Deletion experiments show that the sequences *N*-flanking the conserved Asp is critical for Tcf/Lef binding to *β*-catenin, including G^{13}–T^{20} of human Lef1,[31] and G^{8}–A^{14} of Tcf4.[11, 29-30] **Table 2.1** listed the key binding segments of Lef1 (Q9UJU2), Tcf4 (Q9NQB0), E-cadherin (P12830), and APC-R3(P25054). The *N*-terminal flanking amino acids are shown. θ and Φ are the hydrophobic and aromatic residues, respectively. The sequences that bind to the Tcf4 $G^{13}ANDE^{17}$ binding site are shown in the rectangular red box.

The residues in E-cadherin and APC-R3 that correspond to Tcf4 $G^{13}AN^{15}$ do not bond directly with *β*-catenin. The FP competitive inhibition assay was performed to evaluate the importance of the corresponding sequences of Tcf, cadherin, and APC, as shown in **Table 2.2**. Tcf4 peptide (residues 7–51) has an IC_{50} value of 1.09 nM.[28] Under

Table 2.1 Sequences of human Lef1, Tcf4, E-cadherin and APC-R3.

		DXΘΘXΦXXE	
hLef1	13	GDPELCAT**DEMIPFKDE**GD	31
hTcf4	8	GGDDLGAN**DELISFKDE**GE	26
hE-cadherin	820	TDPTAPPY**DSLLVFDYE**GS	838
hAPC-R3	1478	RVQVLPDA**DTLLHFATE**ST	1496

Table 2.2 Inhibitory activities of Tcf, cadherin, and APC sequences.

Peptides	**IC50, nM**
Tcf (7-51)	1.09 ± 0.04
Tcf (7-51, D16A, E17A)	>200,000
Tcf (18-51)	>200,000
E-cadherin (819-873)	94.66 ± 3.90
E-cadherin (819-873, $p^{826}YDS^{829}/A^{826}AAA^{829}$)	943.7 ± 2.9
APC-R3 (1482-1511)	2690 ± 120
APC-R3 (1482-1511, $D^{1484}ADT^{1487}/A^{1484}AAA^{1487}$)	79290 ± 2830

the same assay, Tcf4 peptide (residues 7–51, D16A/E17A)[28] or a truncated Tcf4 peptide (residues 18–51) did not disrupt β-catenin/Tcf4 interactions even at a concentration of > 200 μM. In contrast, the simultaneous mutation of four residues, $P^{826}YDS^{829}$ of human E-cadherin (residues 819–873) into alanine, only leads to a 10-fold increase in the IC_{50} value. The mutation of four residues, $D^{1484}ADT^{1487}$ of human APC-R3 (residues 1477–1519) into alanine, causes a 30-fold increase in the IC_{50} value. The FP saturation binding assay was also performed for E-cadherin (819–873) and E-cadherin mutant (819–873, $P^{826}YDS^{829}/A^{826}AAA^{829}$). The K_d values of E-cadherin and the E-cadherin $P^{826}YDS^{829}/A^{826}AAA^{829}$ mutation are 80.40 ± 1.1 and 832.1 ± 14.4 nM, respectively. As a comparison, the K_d value of human Tcf4 D16A/E17A is > 100 μM, while the K_d value of native human Tcf4 is 4.71 nM.[28] Based on these results, we proposed that targeting the Tcf

$G^{13}ANDE^{17}$ binding site by mimicking the binding mode of Tcf4 $G^{13}ANDE^{17}$ would lead to the generation of selective small-molecule inhibitors specific for *β*-catenin/Tcf interactions.

2.2.2 Peptidomimetic Strategy to Design Small-Molecule Inhibitors

Pockets A, B, and C of the Tcf GANDE binding site were characterized by SiteMap[33] and multiple copy simultaneous search (MCSS).[34] SiteMap uses a GRID-based method to define the binding site, calculate hydrophobicity, and determine the orientations of H-bond donors and acceptors. The SiteMap analysis indicates that the hydrophobic region is mainly located in pocket B, as the yellow region shown in **Figure 2.3a**. The molecular interaction fields (MIFs) of H-bond donors are primarily from the side chains of K508, R469, K435, S473, R474, R515, R612, and the backbone amide of N430. The MIFs of H-bond acceptors were determined by the side chains of N516, E571, H578, S473, and the backbone carbonyl of K508 (**Figure 2.4**).

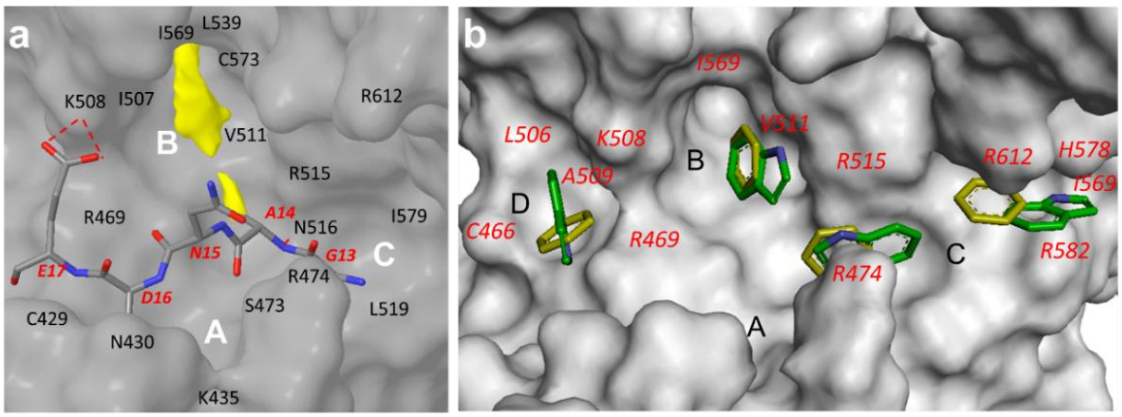


Figure 2.3. Analysis of the Tcf4 $G^{13}ANDE^{17}$ binding site. Pocket A, B, C and D are indicated. **a**. Results of hydrophobic SiteMap analysis. **b**. Representative MCSS-minimized positions of hydrophobic aromatic fragments, benzene (yellow), and indole (green). *β*-Catenin (PDB id, 2GL7) is shown as a space-filled model.

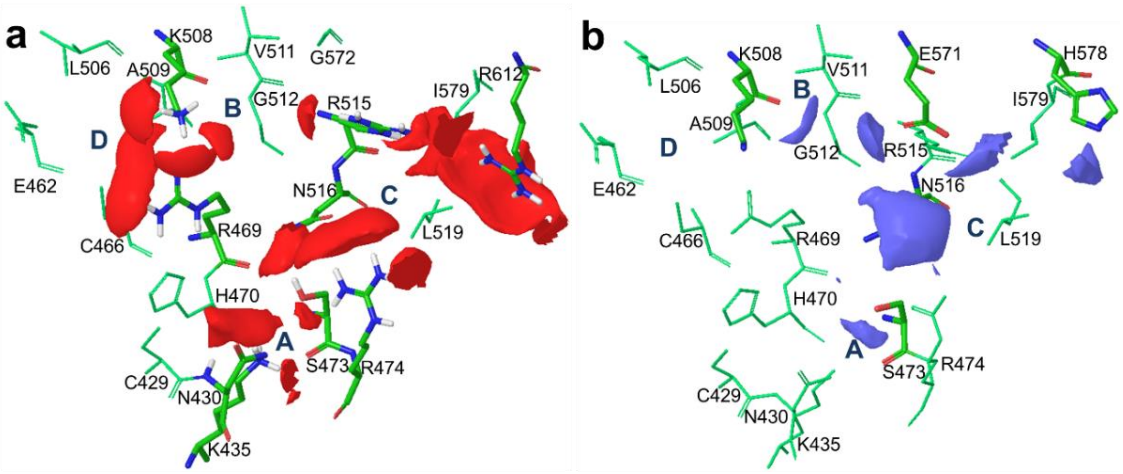


Figure 2.4. Results of H-bonding SiteMap analysis. The *β*-catenin residues are shown green. The residues contribute to the SiteMap contours are colored atom type. **a**. H-bond donor map (red) **b**. H-bond acceptor map (blue).

Ten aromatic, aliphatic, polar and charged functional groups that are commonly found in bioactive small molecules were used in an MCSS analysis (**Table 2.3**). The minima having the second and third most favorable interaction energies with benzene and indole, respectively, located in pocket B (**Figure 2.3b**), where they form hydrophobic interactions with V511 and I569. The minima with the third most favorable interaction energy benzene and the second most favorable interaction energy with indole are located in the arginine channel formed by R474 and R515, where favorable cation–π interactions were observed. In pocket C, benzene (minimum No. 5) and indole (minimum No. 4) form cation–π interactions with R612 and R582 and are stabilized by H578 and I569. In pocket D, benzene (minimum No. 6) and indole (minimum No. 5) are stabilized by R469, A509, C466, and L506.

The MCSS results for the aliphatic apolar fragments, cyclohexane, propane, and isobutane, are described in **Figure 2.5**. The region displaying the most favorable interaction energies with cyclohexane and isobutane are located in pocket B lined by V511, I569,

Table 2.3. Minima in the Tcf4 G[13]ANDE[17] binding site found by the MCSS analysis.

MCSS functional groups	Initial number of copies	Number of Minima	Range of U_{bind}[a] for minima (kcal/mol)		Selected minima	
			From	to	U_{bind}; Pocket	Residues involved
Benzene	2500	19	–1.96	–9.94	No.2: –8.19; B	V511, I569
					No.3: –7.88; C	R474, R515
					No.5: –5.94; C	R612
					No.6: –5.71; D	R469, A509, C466, L506
Indole	5000	187	–3.46	–17.54	No.2: –14.74; C	R474, R515
					No.3: –14.35; B	V511, I569
					No.4: –13.92; C	R612, R582, H578, I579
					No.5: –12.85; D	R469, K508, A509, C466
Cyclohexane	5000	19	–0.01	–4.57	No.1: –4.57; B	K508, V511, I569, G512
					No.2: –3.80; A	K435, H470, R474, C429
					No.3: –2.50; D	A509, K508, C466
					No.4: –2.11; C	L519, N516, R515, S473
					No.6: –2.11; B	N516, R515, G512, R469
Propane	5000	54	–1.54	–8.65	No.1: –8.65; A	K435, H470, C429, S473
					No.2: –8.59; B	K508, V511, I569, G512
					No.4: –7.38; D	A509, K508, C466
					No.6: –7.20; B	N516, R515, S473, R469
Isobutane	5000	49	–1.02	–10.77	No.1: –9.67; B	K508, V511, I569, G512
					No.2: –9.38; A	K435, H470, C429, S473
					No.3: –8.42; B	N516, R515, R469, S473
					No.5: –8.27; D	A509, K508, C466
					No.6: –8.06; C	L519, N516, R515, I579
Dimethyl ether	5000	42	–2.95	–16.08	No.1: –16.08; A	K435
					No.3: –12.89; C	R612
					No.4: –12.84; D	R469
					No.6: –10.71; B	K508
					No.7: –10.68; B	K515
Methanol	5000	47	–2.40	–21.79	No.9: –13.06; C	N516
N-Methylacetamide	5000	108	–1.84	–25.86	No.1: –25.86; A	K435, C429, R474
					No.3: –25.09; D	R469, E462, A509
					No.4: –24.31; C	N516, R474, L519
					No.6: –18.40; C	R612, I579
					No.7: –17.68; B	K508
Acetate anion	5000	21	–10.76	–61.04	No.1: –61.04; A	K435, N430
					No.3: –48.82; D	R469
					No.4: –45.72; B	K508
					No.5: –44.34; C	R474
Methylammonium	5000	20	–0.29	–40.57	No.5: –25.76; B	N516, G512

[a]U_{bind}, the binding energy of a given functional group in each minimized replica obtained from the MCSS calculation, is defined as $U_{bind} = U_{protein–group} + U_{group} − U_{group}^0$, where $U_{protein–group}$ represents the nonbonded interactions between β-catenin and the given functional group. U_{group} represents the internal energy of the functional group within the complex, and U_{group}^0 represents the internal energy of the isolated functional group in vacuum.

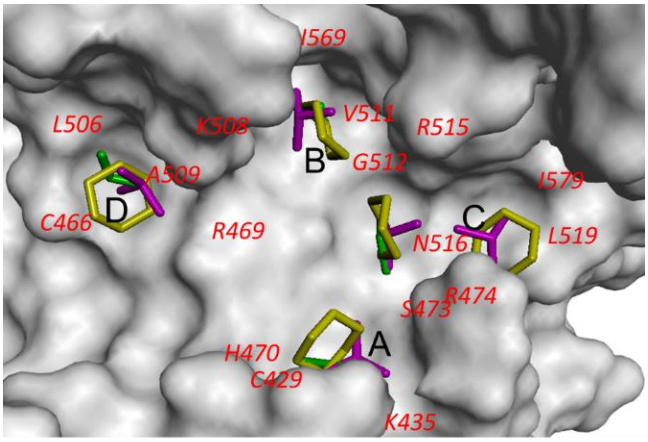


Figure 2.5. Representative MCSS-minimized positions of aliphatic fragments in the Tcf GANDE-binding site of *β*-catenin (PDB id, 2GL7). Cyclohexane, propane, and isobutane are shown yellow, green, and magenta, respectively. *β*-Catenin is shown as a space-filled model. Pockets A, B, C, and D are indicated.

G512, and the side chain carbons of K508. The polar functional groups, dimethyl ether, methanol, and N-methylacetamide, were used to explore the H-bond capacity in these pockets (**Figure 2.6**). The polar functional groups, methanol and N-methyl acetamide, contain H-bond donors. The lowest energy minima of methanol (minimum No. 9) and N-methylacetamide (minim No. 4) as H-bond donors form H-bonds with the side chain carbonyl oxygen of N516, mimicking the H-bond between backbone amide bond of Tcf4 A14 and side chain carbonyl oxygen of *β*-catenin N516.

Based on the results of the above SiteMap and MCSS analyses, a peptidomimetic strategy was employed to design new β-catenin/Tcf inhibitors. The G13ANDE17 sequence of human Tcf4 was used as the starting template. As the previous studies indicated that D16 and E17 of human Tcf4 were critical for Tcf to bind with β-catenin, these two residues were kept in the new inhibitors. The MCSS analysis shows that indole can be located in the R474 and R515 arginine channel. It was merged to the crystallographic binding

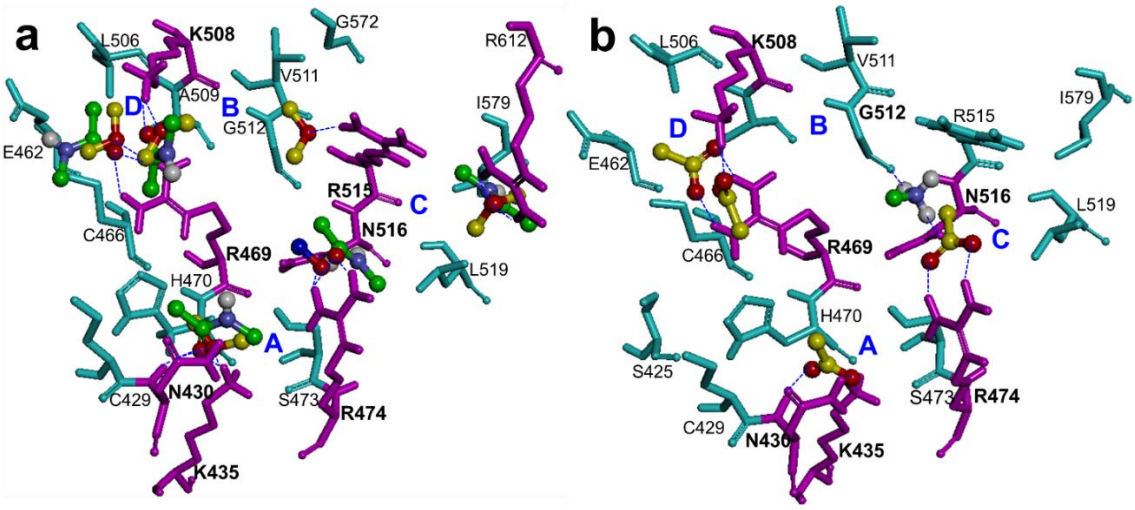


Figure 2.6. Representative MCSS-minimized positions of polar fragments in the Tcf GANDE binding site of *β*-catenin (PDB id, 2GL7). *β*-Catenin is shown as the cyan stick model, and the fragments are shown as the stick–and–ball model. **a**. Polar fragments, dimethyl ether (yellow), methanol (blue), and *N*-methyl acetamide (green) in the active site. **b**. Charged fragments, acetate (yellow), and dimethylammonium (green) in the active site.

conformation of Tcf4 G[13]ANDE[17], which leads to the deletion of residues Gly-Ala. The electron-rich indole ring was used to form cation-π interactions with the arginine channel. The commercially available hydrophobic *D*- and *L*- amino acids identified from SciFinder search were used to replace Asn to explore hydrophobic pocket B. AutoDock 4.2[35] was employed to predict the binding mode of new compounds.

Compound **2.12** in **Figure 2.7a** was selected as the first designed compound for synthesis. The predicted binding mode of **2.12** with *β*-catenin is shown in **Figure 2.7b**. The AutoDock model indicated that the O-methyltyrosine moiety of **2.12** points towards hydrophobic pocket B, while indole ring can form cation-π interactions with R515 (parallel) and R474 (T-shaped). The N-H of indole can form a H-bond with the side chain carbonyl of N516. The indole can also interact with \hydrophobic pocket C. The predicted binding mode of the designed compound matches the proposed key binding elements.

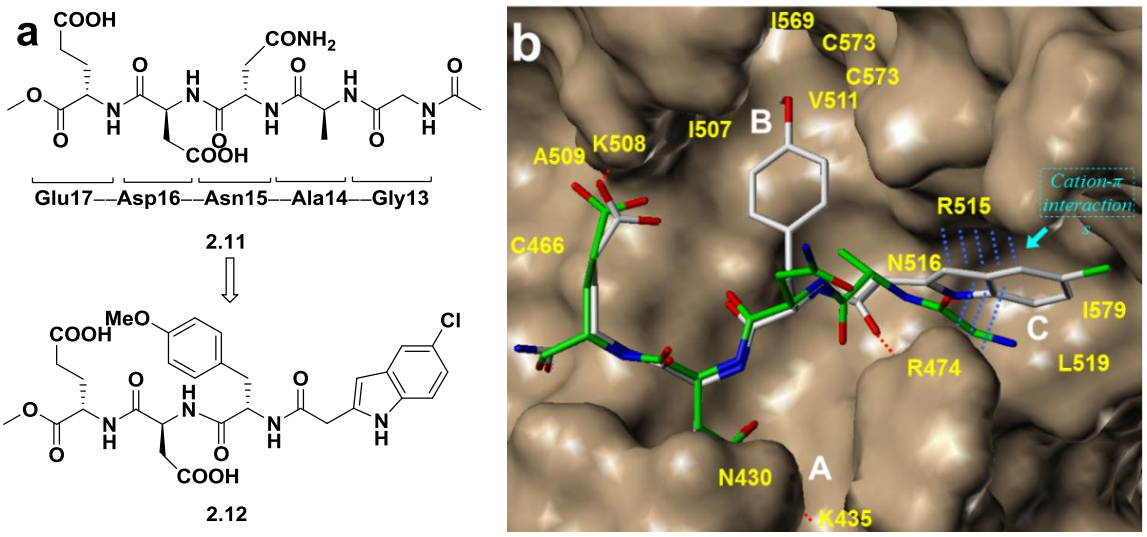


Figure 2.7. Structure-based design of new inhibitors based on Tcf4 G^{13}ANDE^{17}. **a**. Structural evolution from **2.11** to **2.12**. **b**. AutoDock predicted conformation of **2.12** (gray). *β*-Catenin is shown as a space-filled model. Human Tcf4 G^{13}ANDE^{17} (green) is shown as the reference.

Synthesis of compound **2.12** is shown in **Figure 2.8**. The Glu-Asp dipeptide was synthesized through standard peptide coupling reactions (**Figure 2.8a**). The methoxy tyrosine **2.21** was synthesized by a direct alkylation of the phenolic hydroxyl group of Boc-protected tyrosine, followed by the deprotection of the Boc group to yield the desired hydrochloride salts (**Figure 2.8b**). The 2-(5-chloro-1H-indol-2-yl)acetic acid **2.25** was prepared through an intermolecular Wittig reaction using a modified procedure (**Figure 2.8c**).[36] **Figure 2.8d** shows the coupling of the fragments followed by deprotection of all the *tert*-butyl ester and Boc groups in one step to yield compound **2.12**.

Three dipeptides **2.31-2.33** were synthesized as shown in **Figure 2.9** using standard peptide coupling procedures. Together with other intermediates and known inhibitors, they were tested in FP assay to evaluate the contribution of each substructure to inhibitory activity.

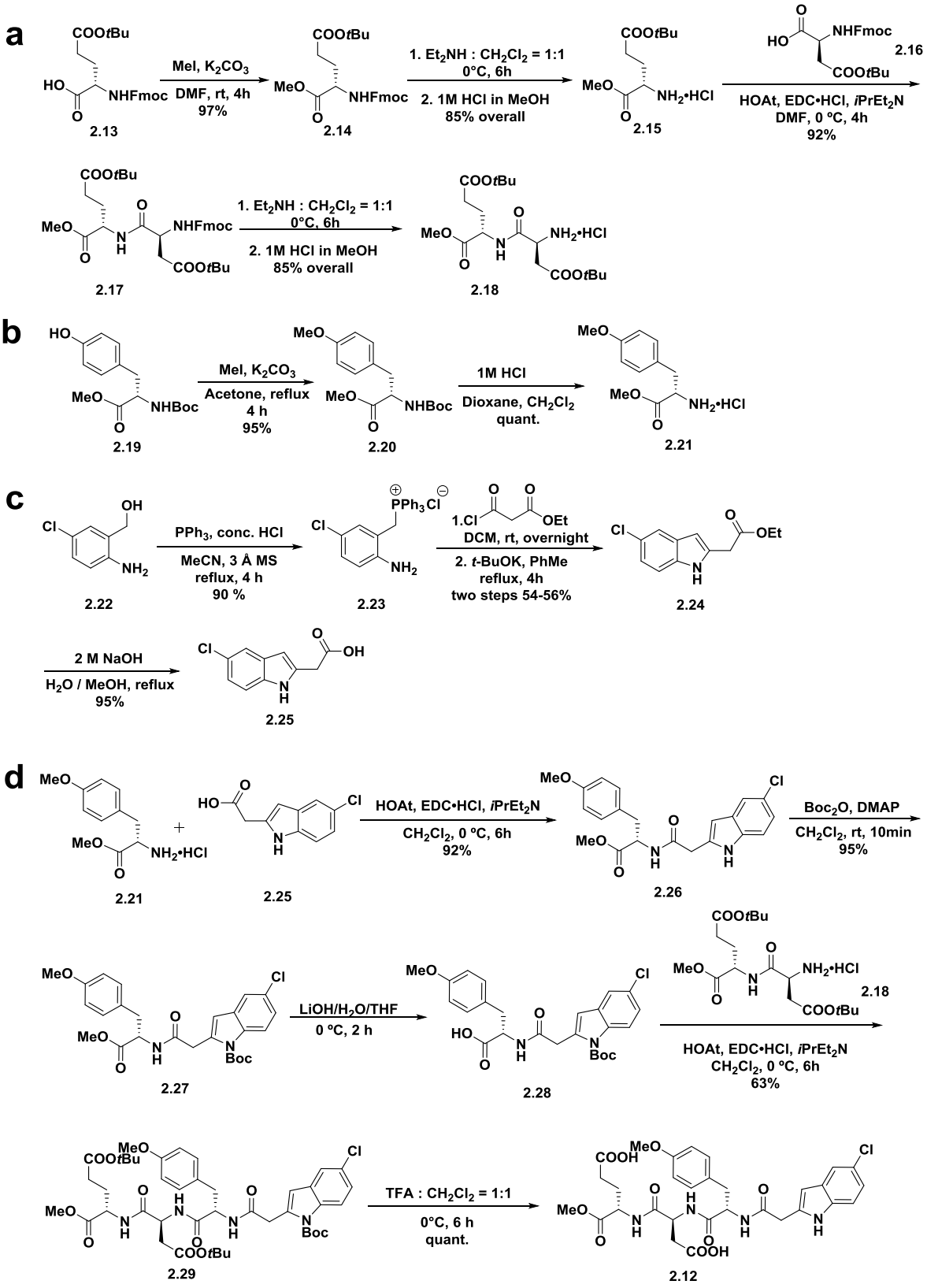


Figure 2.8. Synthesis of **2.12**.

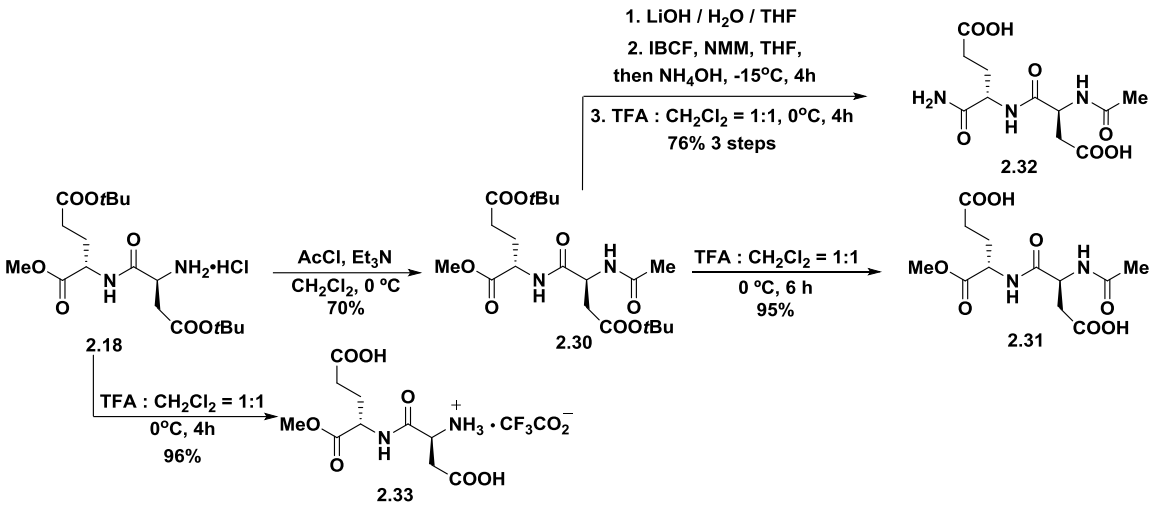


Figure 2.9 Synthesis of dipeptide derivatives.

The FP competitive inhibition assay results of compounds are shown in **Table 2.4**. Compound **2.12** exhibits a K_i value of 5.74 ± 0.76 μM and can completely disrupt the large and tight β-catenin/Tcf PPIs. Compared to **2.11** (K_i = 290.4 ± 1.3 μM), compound **2.12** is 50-fold more potent, indicating the success of the design. Fragments **2.26** and **2.31-2.33** exhibited some inhibitory activities but much were less potent than **2.12**. The FP assay also demonstrated that **2.12** was more potent than known inhibitors KF115–584, CGP049090, PNU74654, iCRT3, iCRT5, and iCRT14 and comparable to PKF115–310 and UU-01.

Table 2.4. The FP assay results of synthesized inhibitors and known inhibitors.

Entry	K_i ± SD (μM)	Compounds[28]	K_i ± SD (μM)
2.11	290.4 ± 1.3	PKF115-584	17.8 ± 2.1
2.12	5.7 ± 0.8	CGP049090	35.6 ± 4.4
2.31	343.0 ± 2.0	PKF118-310	5.8 ± 0.2
2.32	396.6 ± 2.2	PNU74645	181.0 ± 0.7
2.33	119.7 ± 1.2	iCRT3	364.7 ± 6.3
2.26	140.8 ± 1.9	iCRT5	80.3 ± 4.2
UU-01	3.1 ± 0.5	iCRT14	53.5 ± 5.2

2.2.3 Exploration of the Arginine Channel and Pocket B

Additional structural modification involved the exploration of the structure-activity relationship (SAR) for the indole moiety of **2.12**. Compounds **2.34**-**2.40** were designed (**Figure 2.10**).

Compounds **2.34**-**2.35** were synthesized using a method that was similar to the synthesis of **2.12** as shown in **Figure 2.11**. They were designed to study the effects of the substitutions on the indole ring. Compounds **2.36**-**2.40** were synthesized using standard peptide coupling procedures (**Figure 2.12** and **2.13**) from commercially available starting materials.

The FP competitive inhibition assay results are shown in **Figure 2.14**. The K_i of **2.12** is lower than those of **2.34** and **2.35**, indicating chlorine is a better substituent. The importance of the indole moiety of **2.12** was further evaluated by the synthesis of 3,4-disubstituted benzacetamide derivatives, **2.36** and **2.37**. The AutoDock study shows that **2.36** and **2.37** do not have an -NH to form an H-bond with the side chain carbonyl of N516 and have weaker cation-π interactions with R474 and R515 due to fewer π-electrons. In accordance with our expectation, the K_i values of **2.36** and **2.37** were over 60-fold higher than **2.12**, indicating the importance of the H-bond and cation–π interactions for binding. Compounds **2.38**–**2.40** were used to investigate the optimum connectivity for the indole ring. In good agreement with the docking results, substitutions at 2-position of the indole ring (**2.34** and **2.38**) yielded a better inhibitor than those at 3-position (**2.39** and **2.40**). Compared to **2.12**, a relatively flexible 2-substituted side chain, as in **2.34**, can significantly increase the inhibitory activities because the indole can stay comfortably in the arginine channel, and the N-H of the indole ring can form a desired H-bond with N516.

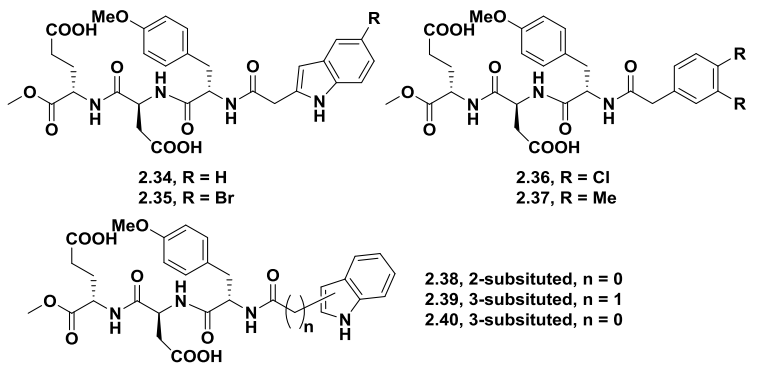


Figure 2.10. Compounds to explore Pocket C.

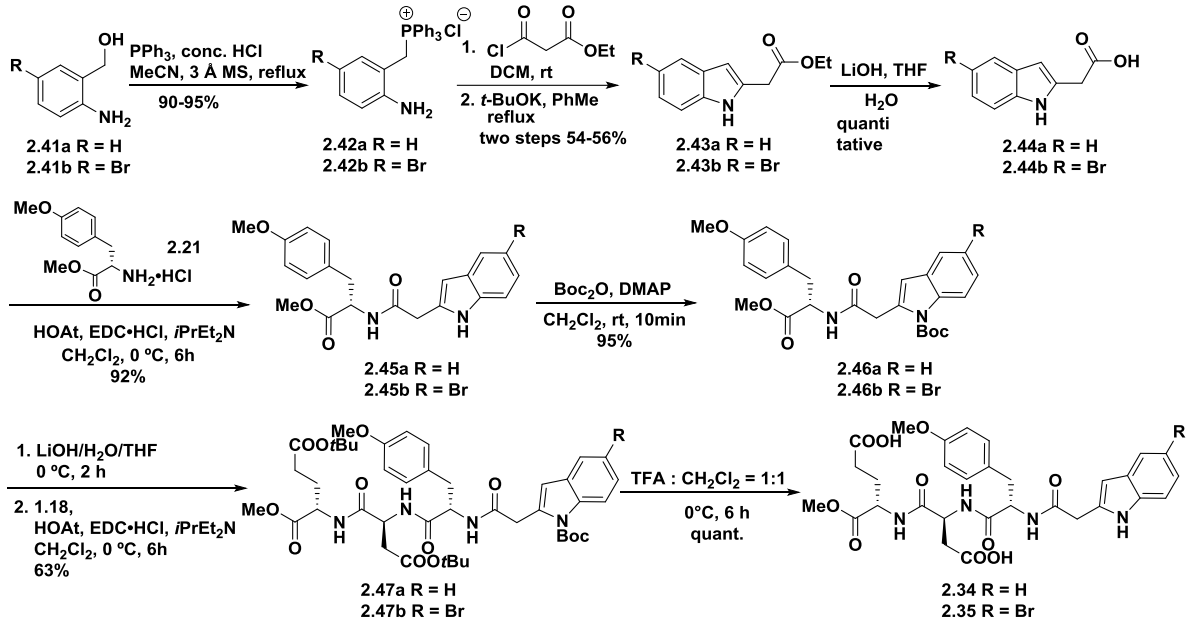


Figure 2.11. Synthesis of compound **2.34**-**2.35**.

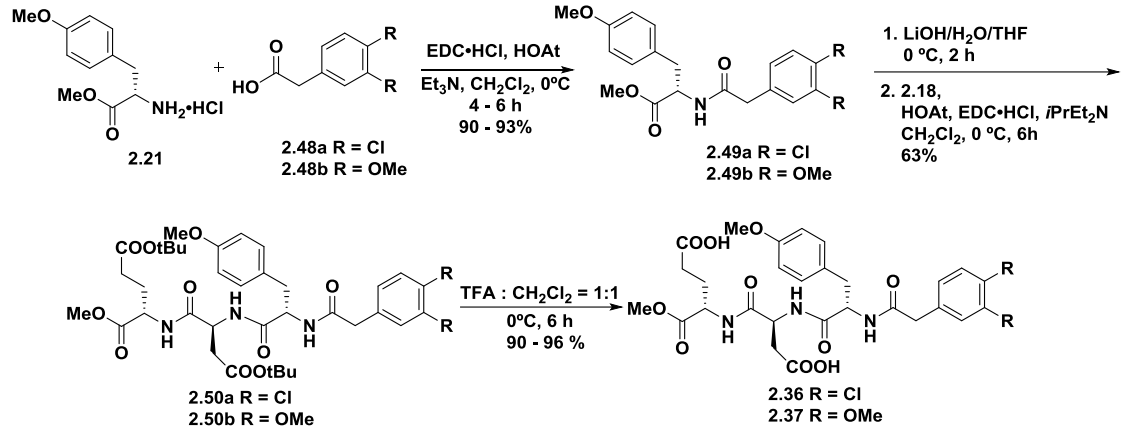


Figure 2.12. Synthesis of compound **2.36**-**2.37**.

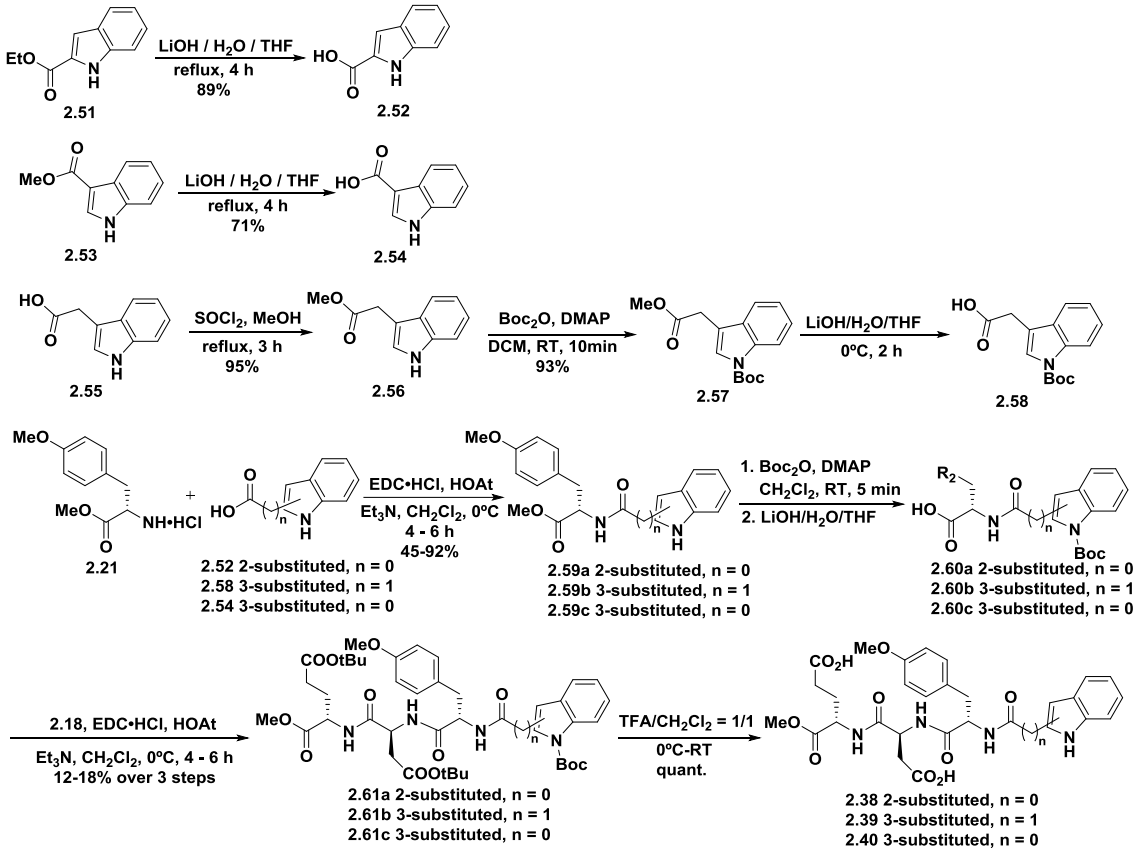
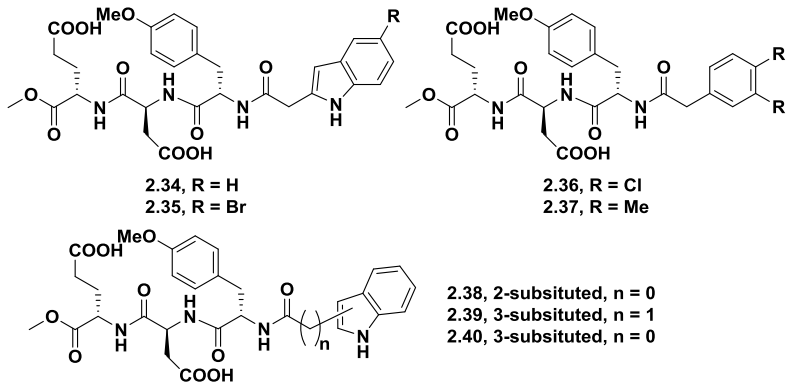


Figure 2.13. Synthesis of compound **2.38**-**2.40**.



Entry	$K_i \pm$ SD (μM)
2.34	45.5 ± 1.2
2.35	13.2 ± 1.2
2.36	333.5 ± 1.1
2.37	305.1 ± 1.9
2.38	51.4 ± 1.1
2.39	60.9 ± 1.1
2.40	301.4 ± 3.0

Figure 2.14. FP competitive inhibition assay results of **2.34-2.40**.

The deep and narrow hydrophobic pocket B was explored using substituted tyrosine derivatives and the other unnatural amino acids. Compounds **2.62-2.67** were designed and synthesized (**Figure 2.15**). The commercially available fragments were prepared and coupled with the 2-indole acetic acid and dipeptide as shown in **Figure 2.16**.

The FP assay results of **2.62-2.67** are shown in **Figure 2.17**. Compound **2.62** has the lowest inhibitory activities in this series presumably coming from the over-sized isopropyl group for pocket B. The phenyl group of **2.63** and the cyclohexyl group of **2.64** only stay at the entrance of pocket B, exhibiting higher K_i values than **2.12**. On the other hand, as expected, more hydrophobic derivatives (compound **2.65-2.67**) have better inhibitory activities. Among this series, **2.66** exhibits the highest inhibitory activity for β-catenin/Tcf PPIs (K_i = 1.36 ± 0.12 μM). It is over 210-fold more potent than **2.11** and represents one of the most potent inhibitors discovered to date.

The activities of the inhibitors were also evaluated using AlphaScreen assay[37] by Dr. Min Zhang (**Table 2.5**). The K_i of **2.66** in the AlphaScreen assay is 1.32 ± 0.56 μM, which more potent than known inhibitors CGP049090 and PKF118–310 (K_i determined in the parallel assays).

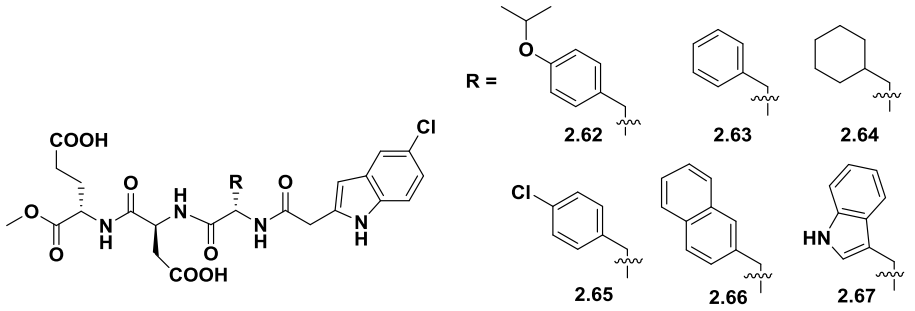


Figure 2.15. Compounds to explore pocket B.

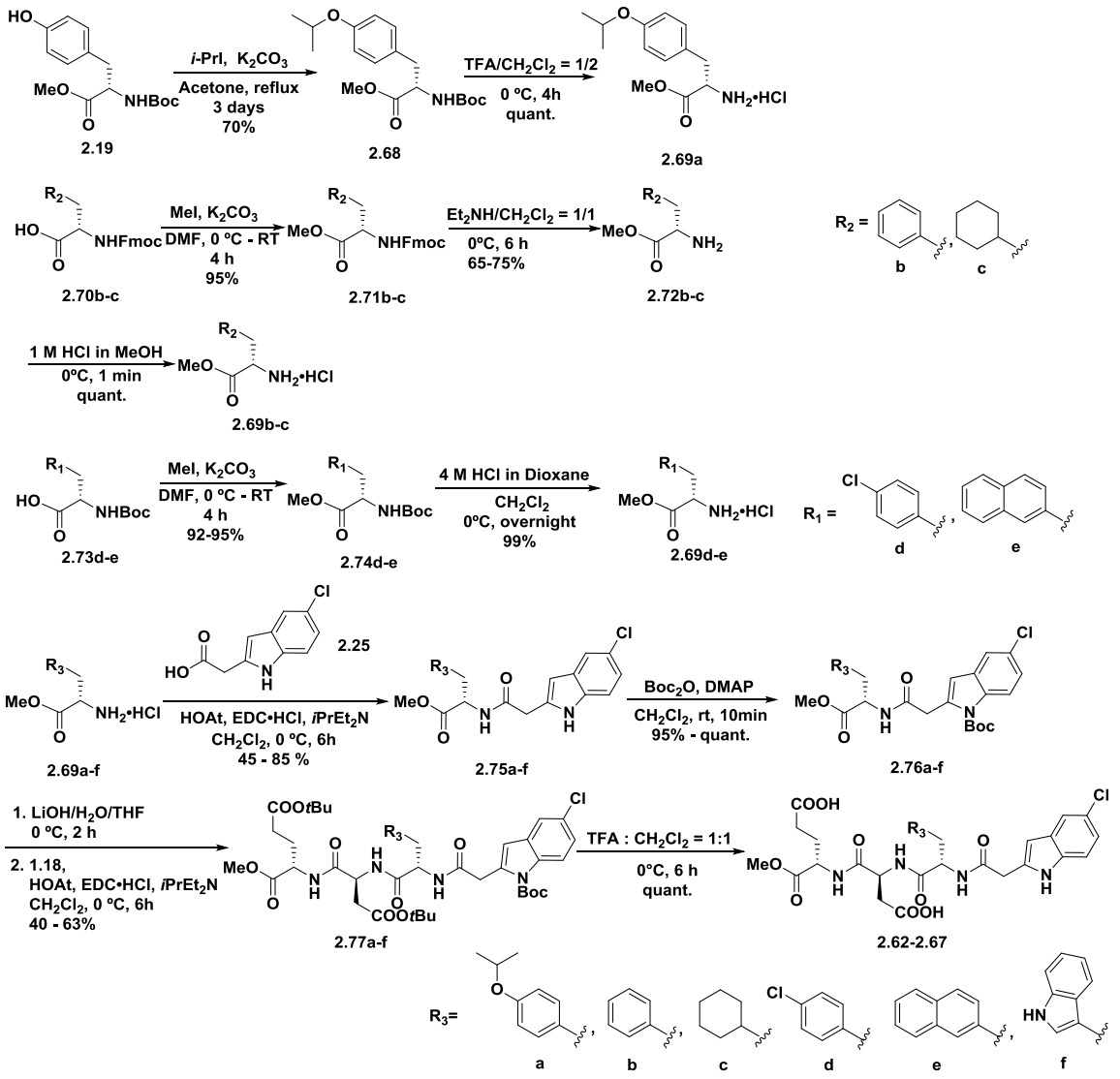
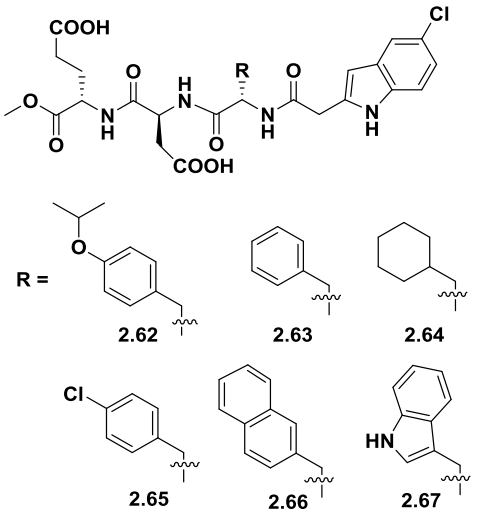


Figure 2.16. Synthesis of compound **2.62**-**2.67**.



Entry	K_i ± SD (μM)
2.62	16.7 ± 1.1
2.63	11.9 ± 0.9
2.64	10.6 ± 0.7
2.65	4.3 ± 0.6
2.66	1.4 ± 0.1
2.67	6.5 ± 1.3

Figure 2.17. FP competitive inhibition assay results of **2.62-2.67**.

Table 2.5. AlphaScreen K_i values of **2.12**, **2.65-2.67** and two known inhibitors.

Entry	K_i ± SD (μM)	Compounds	K_i ± SD (μM)
2.12	4.9 ± 0.8	CGP049090	5.9 ± 0.9
2.65	3.7 ± 0.6	PKF118-310	3.6 ± 0.7
2.66	1.3 ± 0.6	iCRT14	37.0 ± 0.8
2.67	10.9 ± 0.9		

2.2.4 Binding Mode Verification Using Isothermal Titration Calorimetry (ITC) and Site-Directed Mutagenesis Studies

Figure 2.18 shows the AutoDock predicted binding mode of **2.66** with *β*-catenin. Firstly, to verify the inhibitor binds to *β*-catenin instead of Tcf, ITC experiments were performed. **Figure 2.19** confirmed this **2.66** binds to wild-type *β*-catenin with a K_i value of 0.418 μM, but not Tcf (**Figure 2.19**).

The binding mode of **2.66** with *β*-catenin was further evaluated by site-directed mutagenesis studies. Mutated proteins (hydrophobic pocket B: V511S and V511S/I569S; arginine channel pocket C: R474A, R515A and R474A/R515A) were made in the Ji laboratory by Dr. Min Zhang and were verified not affecting *β*-catenin/Tcf binding.[38] The K_i values of **2.12**, **1.65** and **2.66** for wild-type *β*-catenin/Tcf and mutant *β*-catenin/Tcf interactions were determined by FP assays (**Table 2.6**).

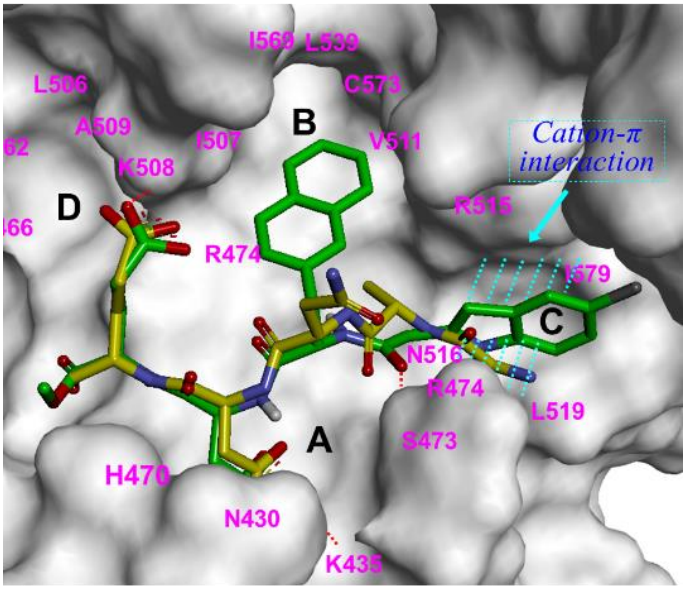


Figure 2.18. Predicted binding mode of **2.66**.

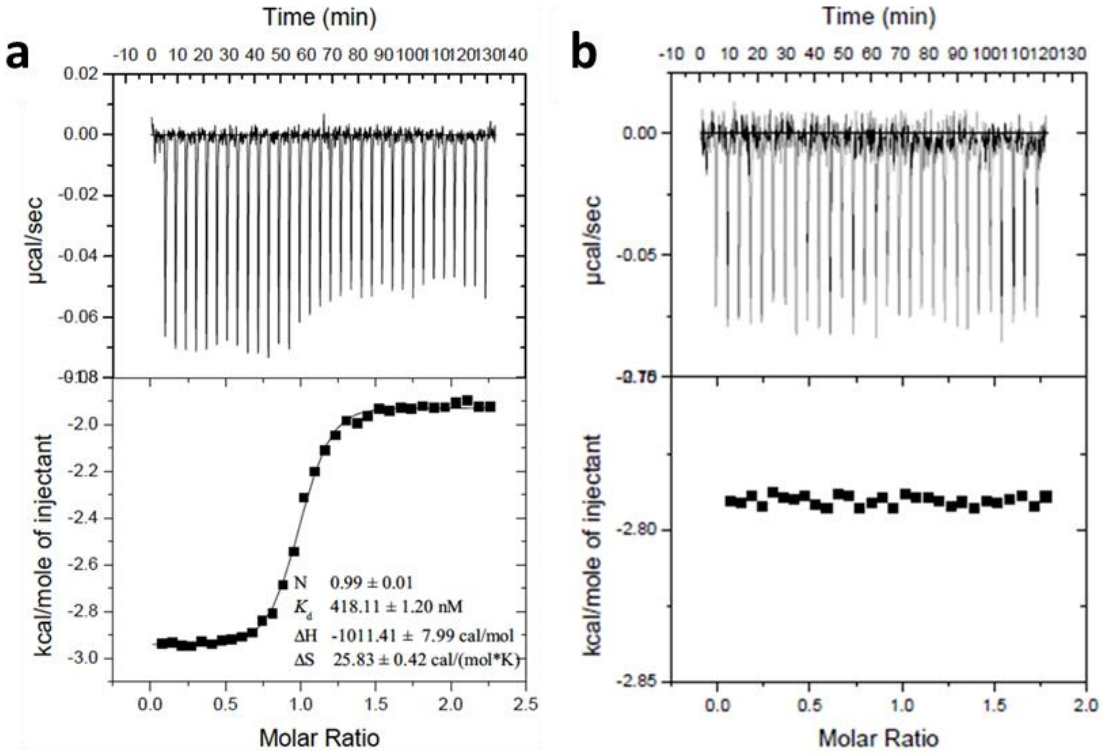


Figure 2.19. ITC studies to determine the binding affinity of compound **2.66** with human *β*-catenin (residues 138–686) and human Tcf (residues 7–51). **a**. *β*-Catenin/**2.66** binding experiment. **b**. Tcf /**2.66** binding experiment.

Table 2.6. Site-directed mutagenesis studies and the FP assay to evaluate the binding mode of **2.12**, **2.65**, and **2.66**.

	K_i, µM		
β-catenin	**2.12**	**2.65**	**2.66**
Wild type	5.74 ± 0.76	4.30 ± 0.59	1.32 ± 0.56
V511S	32.47 ± 2.08	25.98 ± 0.80	5.68 ± 0.78
V511S/I569S	135.2 ± 2.1	83.69 ± 1.64	36.72 ± 0.77
R474A	44.00 ± 1.80	23.28 ± 1.71	11.05 ± 1.24
R515A	35.94 ± 1.04	37.30 ± 0.90	4.69 ± 0.75
R474A/R515A	240.6 ± 2.2	145.8 ± 1.8	52.14 ± 1.75

The mutants in pocket B turned the hydrophobic pocket to a hydrophilic one, which resulted in a more than 20-fold decrease in K_i values for all three compounds, confirming that the hydrophobic interactions were essential for the small molecules to bind.

R474 and R515 of the arginine channel were mutated to alanine. The deletion of two guanidino groups of R474 and R5151 leads to a drastic reduction of the inhibitory activity (greater than 50-fold change for **2.12** and **2.66**), suggesting that the inhibitors indeed interact with the guanidino groups to bind to the protein.

Additionally, we tried to evaluate the binding affinities of **2.66** with mutant *β*-catenin proteins with the ITC studies, but unfortunately, the heat generation was too low to be accurately determined, which insinuated that the inhibitor lost its binding affinity with mutant proteins. Therefore, with the result from above site-directed mutagenesis studies, we are confident that the designed small molecule inhibitors are utilizing the interactions as proposed.

2.2.5 Inhibitor Selectivity Studies

The selectivity assay was established by Dr. Min Zhang using FP to quantify the inhibitor selectivity between *β*-catenin/Tcf, *β*-catenin/E-cadherin, and *β*-catenin/APC interactions.[26] The results are shown in **Table 2.7**. The selectivities of **2.66** for *β*-catenin/Tcf over *β*-catenin/E-cadherin interactions and *β*-catenin/APC interactions are more than 175- and 63-fold, respectively. To our best knowledge, compound **2.66** is the most selective inhibitor reported to this point.

Table 2.7 The FP selectivity assay to determine inhibitor selectivity of the potent small–molecule β-catenin/Tcf inhibitors.

Compounds	$K_i \pm$ SD (μM)			Selectivity	
	β-catenin/ Tcf4	β-catenin/ E-cadherin	β-catenin/ APC-R3	Tcf/ cadherin	Tcf/APC
PKF115–584	17.78 ± 2.13	13.04 ± 0.31	53.85 ± 0.08	0.7	3.0
CGP049090	35.56 ± 4.40	14.39 ± 0.32	62.49 ± 0.05	0.4	1.8
PKF118–310	5.82 ± 0.24	13.30 ± 0.11	171.3 ± 0.1	2.3	29.4
UU-01	3.14 ± 0.48	103.6 ± 0.4	176.0 ± 0.1	33.0	56.0
2.12	5.74 ± 0.76	136.0 ± 4.2	77.21 ± 1.75	23.7	13.5
2.65	4.30 ± 0.59	158.0 ± 2.3	115.0 ± 2.9	36.8	26.7
2.66	1.36 ± 0.12	238.1 ± 2.3	86.63 ± 0.64	175.1	63.7

2.2.6 Cell-based Studies

All cell-based studies were designed and performed in the Ji laboratory by Dr. Min Zhang. The Wnt-responsive luciferase reporter assay was performed using two cell lines, Wnt-activated human embryonic kidney cell (HEK) 293 and colorectal cancer cell SW480. HEK293 cells were co-transfected with *β*-catenin plasmid and TOPFlash (luciferase reporter gene with wild-type Tcf4 binding sites) or FOPFlash (luciferase reporter gene with mutant Tcf4 binding sites). SW480 cells were only transfected with TOPFlash or FOPFlash because this cell type itself harbors deletion APC and has highly elevated *β*-catenin. A known inhibitor of the canonical Wnt signaling pathway, quercetin (**Figure 2.20**),[39] was used as a reference in the cell-based assays.

Compound **2.78** (**Figure 2.21**), an ethyl ester derivative of **2.66**, was synthesized for the cell-based assays since it was hard for **2.66** to pass through the cell membrane with two carboxylic acid groups. To our delight, compound **2.78** exhibited better inhibitory activity than **2.66** with IC_{50} values of 28.73 ± 3.09 and 37.57 ± 3.23 μM in luciferase reporter assays for SW480 and Wnt-activated HEK293 cells, while not reducing the FOPFlash luciferase activities (**Figure 2.22**), which is consistent with the specificity of this inhibition.

The downregulation of **2.78** to the Wnt signaling pathway was evaluated by real-time PCR study (**Figure 2.23**). *AXIN2* and *LGR5* are the specific targets of the canonical Wnt signaling pathway. *Cyclin D1* and *c-myc* are two important target genes of canonical Wnt signaling that promote tumorigenesis. Compound **2.78** down-regulated the transcription of *AXIN2*, *LGR5*, *cyclin D1*, and *c-myc* in dose-dependent manners in SW480 cells.

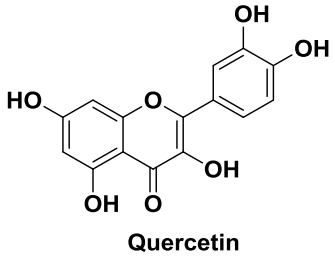
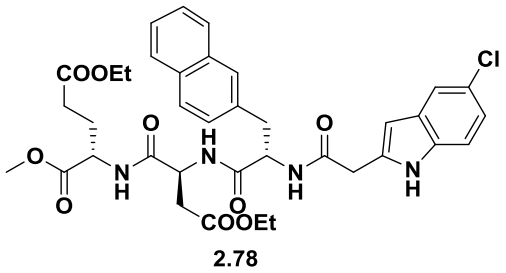


Figure 2.20. Structure of quercetin.



Entry	IC_{50}, μM	
	SW480	**HEK293**
2.66	231.9 ± 3.2	n.d.
2.78	28.7 ± 3.1	37.6 ± 3.2
quercetin	66.7 ± 3.6	69.5 ± 3.3

Figure 2.21. Wnt-responsive luciferase reporter assays for **2.66**, **2.78** and quercetin.

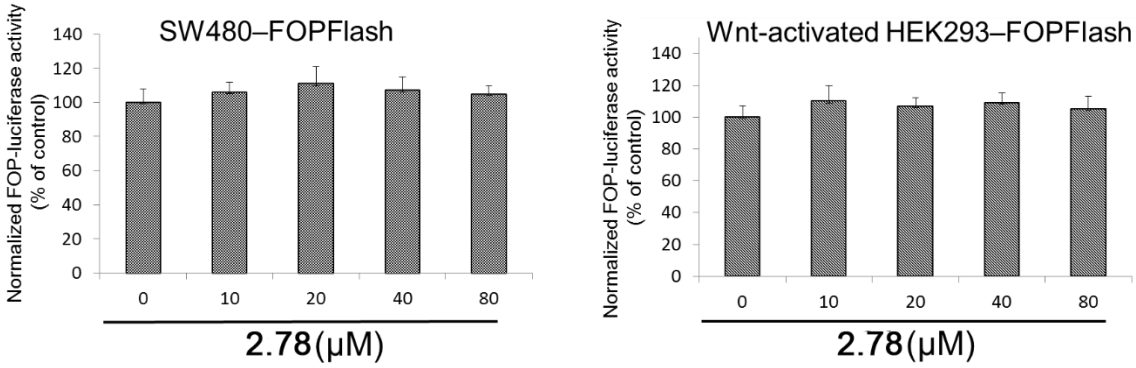


Figure 2.22. FOPFlash luciferase reporter assay results of **2.78**.

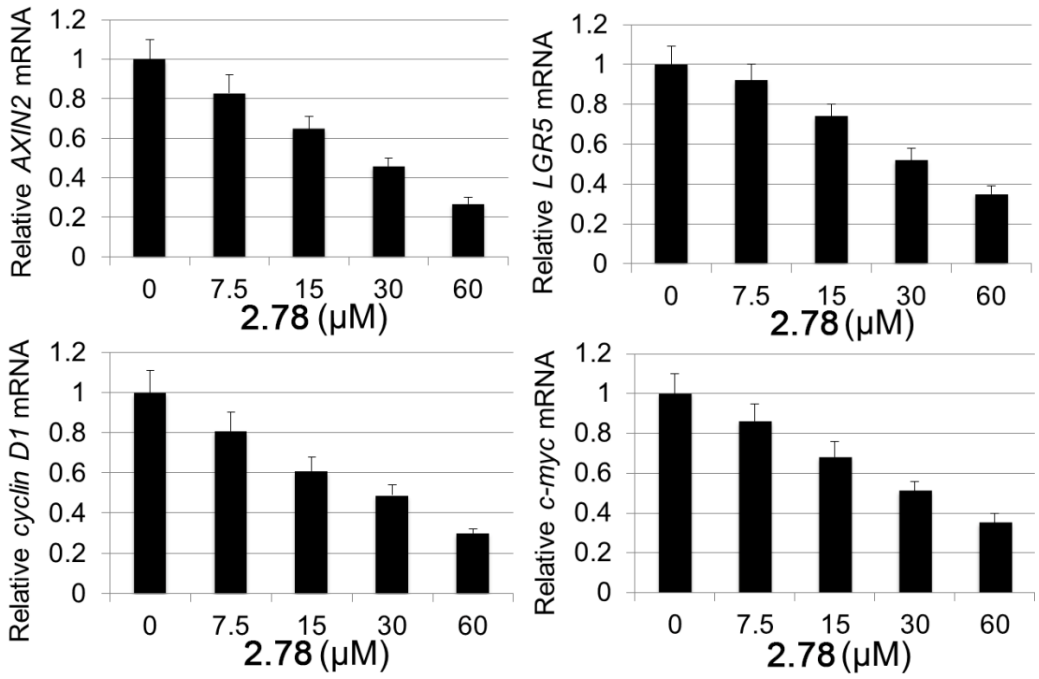


Figure 2.23. Quantitative real-time PCR study to determine the changes in mRNA expression of *AXIN2*, *LGR5*, *cyclin D1*, and *c-myc* in response to different concentrations of **2.78**. SW480 cells were used.

The protein expression levels of *cyclin D1*, *c-myc*, and *β*-catenin in SW480 cells were also examined by the Western blot analysis (**Figure 2.24**). The protein expression levels of *cyclin D1* and *c-myc* were significantly reduced after the treatment of **2.78**. The protein expression level of *β*-catenin was not affected by the treatment of **2.78**, indicating that **2.78** does not inhibit the upstream sites of canonical Wnt signaling.

Co-immunoprecipitation experiments were performed to evaluate inhibitor selectivity in colorectal cancer cell HCT116. As shown in **Figure 2.25**, compound **2.78** inhibited *β*-catenin/Tcf PPIs in a dose-dependent manner but had no effect on *β*-catenin/E-cadherin and *β*-catenin/APC PPIs at the highest examined concentration.

Compound **2.78** also inhibited the growth of colorectal cancer cells with the IC_{50} values of 10.77 ± 1.18, 11.01 ± 1.15, and 28.83 ± 1.26 μM for colorectal cancer cells

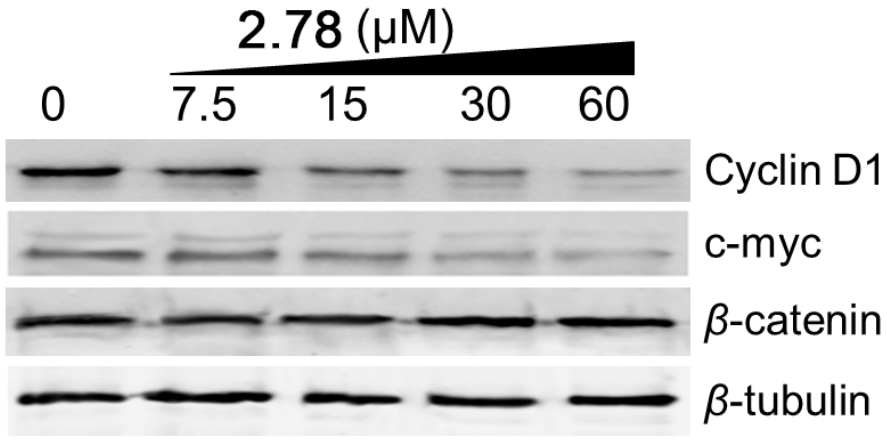


Figure 2.24. Western blot analysis to monitor the changes of protein expression of *c-myc*, *cyclin D1*, and *β*-catenin in response to different concentrations of **2.78**. *β*-Tublin was used as an internal reference, and SW480 cells were used.

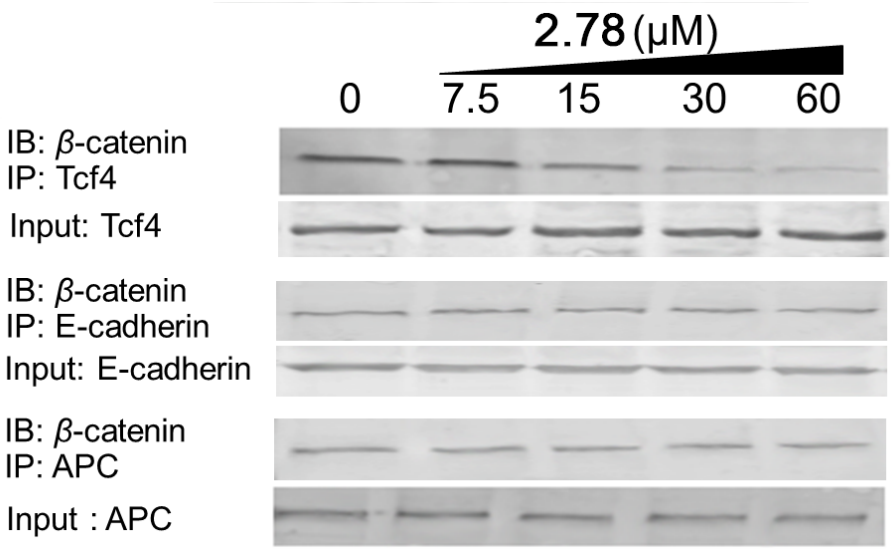


Figure 2.25. Co-immunoprecipitation experiments to evaluate inhibitor selectivity of **2.78** for *β*-catenin/Tcf over *β*-catenin/E-cadherin and *β*-catenin/APC interactions. IP, immunoprecipitation; IB, immunoblotting; input, 10% amount of cell lysate.

SW480, HT29, and HCT116, respectively (**Table 2.8**). This compound exhibited a much higher IC_{50} value of 84.02 ± 3.62 μM for normal cell HEK293, again demonstrating the selectivity of this inhibitor.

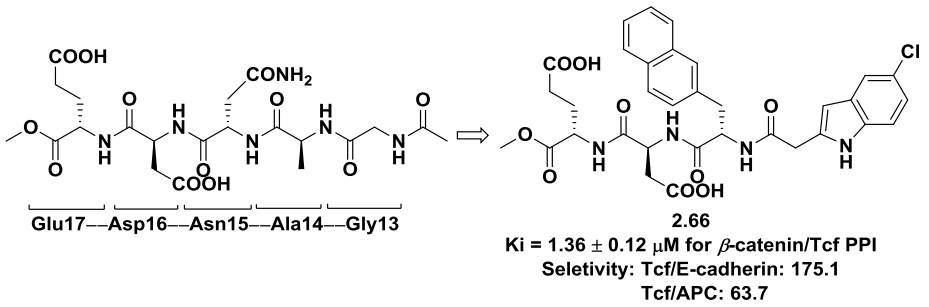
2.3 Conclusion

The discovery of small-molecule inhibitors that selectively disrupt one protein-protein interface not only provides a class of new therapeutic targets but also greatly facilitates our understanding of the structure and dynamic aspects of cellular networks. In this chapter, alanine scanning and the FP assays were used to identify a selective binding site that is crucial for *β*-catenin/Tcf interactions but much less important for *β*-catenin/cadherin and *β*-catenin/APC interactions. A new peptidomimetic strategy incorporating SiteMap and MCSS was used to explore new pockets/new binding elements adjacent to the hot spot. Hydrophobic pocket B and the positively charged arginine channel (pocket C) were employed to design new potent and selective inhibitors. After three rounds of structural optimization, compound **2.66** was identified as the best hit (**Figure 2.26**). Site-directed mutagenesis and structure−activity relationship studies were performed to evaluate the binding mode of **2.66**.

Compound **2.66** has a K_d value of 0.418 μM for binding to *β*-catenin and a K_i value of 1.36 μM to completely disrupt *β*-catenin/Tcf interactions. It also exhibits dual selectivity for *β*-catenin/Tcf over *β*-catenin/E-cadherin and *β*-catenin/APC interactions. Compound **2.78**, an ethyl ester derivative of **2.66**, can effectively pass the cell membrane and inhibit canonical Wnt signaling and the growth of colorectal cancer cells. This compound provides an excellent starting point to generating more potent and selective inhibitors specific for *β*-

Table 2.8 MTs assay to monitor the effects of **2.66**, **2.78**, and quercetin on the growth of colorectal cancer cells. n.d., not determined.

Entry	IC_{50}, μM			
	SW480	**HT29**	**HCT116**	**HEK293**
2.66	152.6 ± 3.7	130.4 ± 4.9	87.30 ± 6.07	n.d.
2.78	10.77 ± 1.18	11.01 ± 1.15	28.83 ± 1.26	84.02 ± 3.62
quercetin	60.09 ± 4.75	60.03 ± 2.93	54.31 ± 3,28	n.d.



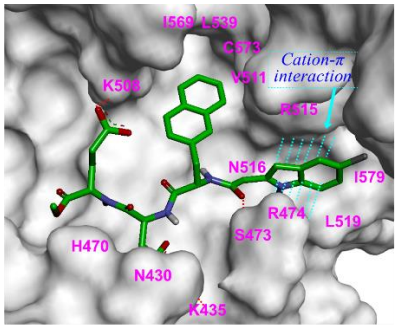


Figure 2.26. From peptide to small molecule **2.66**.

catenin/Tcf interactions. Furthermore, we envision that further modification of **2.66** with a focus on pockets B and C will lead to new inhibitors with better selectivity along with higher inhibitor potency.

2.4 Methods and Experiments

2.4.1 Protein Structure for Computer Modeling

The crystallographic coordinates for human *β*-catenin (PDB id, 2GL7, 2.60 Å resolution, R_{cryst} = 0.223) were obtained from the Research Collaboratory for Structural Bioinformatics (RCSB) protein database. All computational work was performed on Linux Red Hat 6.2 Workstations. The preparation of the crystal structure and molecular modeling were achieved with the commercially available Accelrys Discovery Studio 3.0 (http://accelrys.com/), Schrodinger (http://www.schrodinger.com/) and SYBYL-X 2.0 (http://www.tripos.com) software packages. The missing side chains of *β*-catenin were added in SYBYL X2.0. The protonation states of the residues were set to pH 7.0 when adding the hydrogens. The AMBER 7 force field 99 within SYBYL X2.0 was used to optimize the orientation of hydrogen atoms and the missing side chains of the protein and structural waters. After the protein structure was optimized. Chains B (Tcf4), C (BCL9), D (the second monomer of *β*-catenin), E (the second monomer of Tcf4), F (the second monomer of BCL9), and solvent molecules were removed, leaving only one monomer of *β*-catenin for further calculation. The residues in the Tcf4 $G^{13}ANDE^{17}$ binding site of *β*-catenin include G422, S425–N426, T428–N430, K435, E462–P463, I465–C466, R469–H470, S473–R474, Q482, L506–A509, V511–G512, R515–N516, L519, L536, L539, R565, E568–I569, E571–C573, G575–A576, H578–I579, R582, N609, R612–V613, E620,

and Y654.

2.4.2 Fragment Docking using MCSS

The functional groups chosen for the MCSS analysis were benzene, indole, cyclohexane, propane, isobutane, dimethyl ether, methanol, *trans*-*N*-methyl-acetamide, acetate anion, and methylammonium cation. The binding site area in the MCSS simulation was defined to include all the residues of the Tcf4 $G^{13}ANDE^{17}$ binding site. 2500–5000 Replicas of a given functional group were randomly distributed inside the binding site (The distance threshold between fragment and protein was set to 0.9 Å) and then simultaneously and independently energy minimized. A distance-dependent dielectric model was used to approximate the solvent. Default iteration profile and energy thresholds were used. Pairs of molecules were considered to be identical if the rmsd between them was less than 0.2 Å. The minima with the binding energy lower than 0 kcal/mol were retrieved in the analysis. All of the calculations were performed using the CHARMM force field and Momany and Rone partial charge.[40] At the end of the MCSS run, energy minima within 2.0 Å rmsd of each other were considered as one cluster, and the copy in each cluster with the lowest energy was selected as cluster representative.

2.4.3 Ligand Docking using AutoDock 4.2[35]

In AutoDock studies, only the polar hydrogen atoms remained on the protein structure, and Kollman united atom charges were assigned. The 3D structures of the ligands were built, and the partial atomic charges were calculated using the Gasteiger–Marsili method. The rotatable bonds in the ligands were defined using AutoTors, which also united

the nonpolar hydrogens and partial atomic charges to the bonded carbon atoms. The grid maps were calculated using AutoGrid. The AutoDock area was defined to include all the residues of the Tcf4 $G^{13}ANDE^{17}$ binding site, and the grid spacing was set to 0.375 Å. Docking was performed using the Lamarckian genetic algorithm, and the pseudo-Solis and Wets method were applied to the local search. Each docking experiment was performed 100 times, yielding 100 docked conformations. Other settings were the standard default parameters. All of the ligands followed the same docking protocol. The results of the docking experiments were evaluated by the auxiliary clustering analysis and a visual inspection to match the proposed binding elements.

2.4.4 FP Competitive Inhibition Assay[26, 37]

Experiments were performed in 96-well Microfluor 2 black plates on a Synergy 2 plate reader (Biotek). The polarization was measured at room temperature with an excitation wavelength at 485 nm and an emission wavelength at 535 nm. The FP experiments were performed in an assay buffer of 137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 2 mM KH_2PO_4, 100 μg/mL of bovine gamma globulin, and 0.01% Triton-X 100. The final reaction volume was set to 100 μL. For the β-catenin/Tcf assay, 10 nM of β-catenin (residues 142–686) was incubated with 2.5 nM of *C*-terminally fluorescein-labeled human Tcf4 (residues 7–51) for 30 min at 4 ℃, and then different concentrations of the tested peptides and compounds in assay buffer were added. The negative control (equiv to 0% inhibition) refers to 2.5 nM Tcf4 fluorescence tracer and 10 nM β-catenin in assay buffer without the tested compound. The positive control (equiv to 100% inhibition) refers to only 2.5 nM Tcf4 fluorescence tracer in assay buffer. For the β-catenin/cadherin

assay, 150 nM human *β*-catenin (residues 138−686) was incubated with 5 nM of *C*-terminally fluorescent-labeled human E-cadherin (residues 819−873) in assay buffer for 30 min at 4 ℃. The negative control refers to 5 nM E-cadherin fluorescence tracer and 150 nM *β*-catenin in assay buffer with no inhibitor presenting. The positive control refers to 5 nM E-cadherin fluorescence tracer in assay buffer. For the *β*-catenin/APC-R3 assay, 2000 nM human *β*-catenin (residues 138−686) was incubated with 5 nM of *C*-terminally fluorescent-labeled human APC-R3 (residues 1477−1519) in assay buffer for 30 min at 4 ℃. The negative control refers to 5 nM APC-R3 fluorescence tracer and 2000 nM *β*-catenin in assay buffer without the tested compound. The positive control refers to 5 nM APC-R3 fluorescence tracer in assay buffer. Each assay plate was covered black and gently mixed on an orbital shaker at 4 ℃ for 2.5 h to reach equilibrium before the polarization values were read. The background of the tested inhibitors was corrected by subtracting the raw intensity values of the sample background well (all components except probe) from the raw intensity values of the corresponding test wells (all components). The IC_{50} values were determined by GraphPad Prism 5.0. The K_i values were derived from the IC_{50} values.[26] All of the experiments were performed in triplicate and carried out in the presence of 1% DMSO for small-molecule inhibitors. The results were expressed as mean ± standard deviation. The Tcf/cadherin selectivity ratios were calculated based on the respective K_i value of *β*-catenin/E-cadherin interactions over that of *β*-catenin/Tcf4 interactions. The Tcf/APC selectivity ratios were calculated based on the respective K_i value of *β*-catenin/APC-R3 interactions over that of *β*-catenin/Tcf4 interactions.

2.4.5 AlphaScreen Assay

Experiments were performed in white opaque 384-well plates from PerkinElmer (Waltham, MA), and the samples were read on a Synergy 2 plate reader (Biotek, Winooski, VT) with a sensitivity setting of 200 using AlphaScreen protocol with excitation at 680 nm and emission at 570 nm. All dilutions were made in 1 × assay buffer containing 25 mM HEPES (pH 7.4), 100 mM NaCl and 0.1% BSA to minimize nonspecific interactions. In the AlphaScreen competitive inhibition assay, 20 nM of *N*-terminal His_6-tagged β-catenin (residues 142-686) was incubated with 5 nM of *C*-terminal biotinylated Tcf4 45-mer (residues 7-51) for 30 min at 4 ℃, and then different concentrations of the tested inhibitors in the assay buffer were added to make a final volume of 20 μL. The mixture was incubated at 4 °C for 2 h. Donor and acceptor beads were added to a final concentration of 10 μg/mL in 25 μL of assay buffer. The mixture was incubated at 4 °C for 1 h. The IC_{50} value was determined by nonlinear least-square analysis of GraphPad Prism 5.0. The K_i values were derived from the IC_{50} values by a reported method.[41] Experiments were performed in triplicate and carried out in the presence of 1 % DMSO.

2.4.6 ITC Experiments

ITC measurements were performed at 30 ℃ using a VP-ITC (Microcal, GE Healthcare Life Sciences). The concentrations of **2.66** and β-catenin were set to 80 μM and 8 μM, respectively, in a buffer of 20 mM HEPES, pH 7.5, 200 mM NaCl, 1mM TCEP, 0.01% Triton X-100, and 8% glycerol. Each titration experiment was initiated by a 2 μL injection, followed by 30–35 injections with 8 μL per injection. Blank titrations, which were carried out by injecting the compound into the buffer, were subtracted from each data

set. The association constant K_A, enthalpy change (ΔH), and stoichiometry N were obtained from fitting the data using the Origin software package (Microcal). The K_d, the free energy change ΔG, and the entropy change ΔS were obtained from the basic thermodynamic relationships $K_d = K_A^{-1}$, $\Delta G = -RT\ln K_A$, and $\Delta G = \Delta H - T\Delta S$.

2.4.7 Site-Directed Mutagenesis Studies

β-Catenin mutants V511S, V511S/I569S, R474A, R515A, and R474A/R515A were generated using the overlapping PCR technique. The template for the mutagenesis reactions was the full-length wild-type *β*-catenin in pET-28a. KOD hot start DNA polymerase (Novagen) was used for all the experiments. Mutants were confirmed by direct sequencing (Core facility, University of Utah). The templates used for *β*-catenin double mutation were V511S for V511S/I569S and R474A for R474A/R515A, respectively. The double mutants were again confirmed by direct sequencing. Following the confirmation of the sequence, *β*-catenin mutants were cloned into a pET-28 vector and transformed into *E. coli* BL21 DE3.

2.4.8 Cell Transfection and Luciferase Assay

FuGENE6 (Promega) 96 well plate format was used for the transfection of HEK293 and SW480 cells according to the manufacturer's instructions. HEK293 cells were cotransfected with 45 ng of TOPFlash or FOPFlash reporter gene, 135 ng pcDNA3.1–*β*-catenin and 20 ng pCMV-RL normalization reporter gene. SW480 cells were cotransfected with 60 ng of TOPFlash or FOPFlash reporter gene and 40 ng pCMV-RL normalization reporter. Cells were cultured in DMEM and 10% FBS at 37 ℃ for 24 h and the different

concentrations of inhibitors or DMSO were then added. After 24 h, the luciferase reporter activity was measured using the Dual-Glo system (Promega). Normalized luciferase activity in response to treatment with small-molecule compounds was compared with that obtained from cells treated with DMSO. Experiments were performed in triplicate.

2.4.9 Quantitative Real-Time PCR Analysis

SW480 cells at 1×10^6/mL were treated with different concentrations of inhibitor for 24 h. Total RNAs were extracted with TRIzol (15596026, Life Technologies), and the cDNA was synthesized with the superscript III first-strand kit (18080-051, Invitrogen). Quantitative PCR (qPCR) was performed using the iQ^{TM} SYBR green supermix kit (170-8880, BIO-RAD) on an iQ^5 multicolor real-time PCR reaction system (BIO-RAD). The threshold cycle (C_T) values were normalized to that of internal reference *GAPDH*. The primer pairs for human *GAPDH* were forward: 5′-GAAGGTGAAGGTCGGAGTC-3′, and reverse: 5′-GAAGATGGTGATGGGATTTC-3′, for human *AXIN2* forward: 5′-AGTGTGAGGTCCACGGAAAC-3′ and reverse: 5′-CTTCACACTGCGATGCATTT-3′, for human *LGR5* forward: 5′-TGCTGGCTGGTGTGGATGCG-3′ and reverse: 5′-GCCAGCAGGGCACAGAGCAA-3′, for human *c-myc* forward: 5′-CTTCTCTCCGTCCTCGGATTCT-3′ and reverse: 5′-GAAGGTGATCCAGACTCTGACCTT-3′, and for human *cyclin D1* forward: 5′-ACAAACAGATCATCCGCAAACAC-3′, and reverse: 5′-TGTTGGGGCTCCTCAGGTTC-3′. Experiments were performed in triplicate.

2.4.10 Western Blotting

SW480 cells at 1 × 10^6 cells/mL were treated with different concentrations of inhibitor for 24 h. Cell were lysed in buffer containing 50 mM Tris (pH 7.4), 150 mM NaCl, 1% Nonidet P-40, 0.5% sodium deoxycholate, 0.1% SDS, and protease inhibitors. After centrifugation at 12,000 rpm for 20 min at 4 ℃, the supernatant was loaded onto an 8% SDS-polyacrylamide gel for electrophoretic analysis. Separated proteins were transferred onto nitrocellulose membranes for immunoblot analysis. IRDye 680LT goat anti-mouse IgG (827-11080, LiCOR) or IRDye 800CW goat anti-rabbit IgG (827-08365, LiCOR) was used as the secondary antibodies. The images were detected by the Odyssey Infrared Imaging System (LiCOR). Experiments were performed in duplicate.

2.4.11 Co-immunoprecipitation Assay

HCT116 cells at 1 × 10^6/mL were treated with different concentrations of inhibitor for 24 h. Cells were lysed in buffer containing 50 mM Tris, pH 7.4, 150 mM NaCl, 1% Nonidet P-40, 2 mM EDTA, and protease inhibitors. The lysates were preadsorbed to A/G plus agarose (sc-2003, Santa Cruz Biotechnology, Inc.) at 4 ℃ for 1 h. Preadsorbed lysates were incubated with a specific primary antibody overnight at 4 ℃. A/G plus agarose was then added to the lysates mixture and incubated for 3 h. The beads were washed 5 times with the lysis buffer at 4 ℃. The bound protein was eluted by boiling in the SDS sample buffer and loaded onto 8% SDS-polyacrylamide gel for electrophoretic analysis. Separated proteins were transferred onto nitrocellulose membranes for immunoblot analysis. IRDye 680LT goat anti-mouse IgG (827-11080, LiCOR) was used as the secondary antibody. The images were detected by the Odyssey Infrared Imaging System (LiCOR). Experiments

were performed in duplicate.

2.4.12 MT's Cell Viability Assay

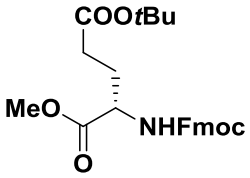
Colorectal cancer cell lines, SW480, HT29, and HCT116, and normal cell HEK293 were seeded in 96 well plates at 4×10^3 cells/well, maintained overnight at 37 ℃, and incubated in the presence of inhibitor at various concentrations. Cell viability was monitored after 72 h using a freshly prepared mixture of 1 part phenazine methosulfate (PMS, Sigma) solution (0.92 mg/mL) and 19 parts 3-(4,5-dimethylthiazol-2-yl)-5-(3-carboxymethoxyphenyl)-2-(4-sulfophenyl)-2H-tetrazolium (MTs, Promega) solution (2 mg/mL). Cells were incubated in 10 μL of this solution at 37 ℃ for 3 h and A_{490} was measured. The effect of each compound is expressed as the concentration required to reduce A_{490} by 50% (IC_{50}) relative to vehicle-treated cells. Experiments were performed in triplicate.

2.4.13 Chemistry

All experiments were conducted under anhydrous conditions in an atmosphere of argon, using flame-dried apparatus and employing standard techniques in handling air-sensitive materials. Dichloromethane (CH_2Cl_2), acetonitrile (CH_3CN), tetrahydrofuran (THF), dimethylformamide (DMF) were degassed with nitrogen and passed through JC Meyer solvent systems. All reagents were purchased from Sigma-Aldrich and used as received. Aqueous solutions of sodium bicarbonate ($NaHCO_3$), sodium chloride (brine), and ammonium chloride (NH_4Cl) were saturated. Analytical thin layer chromatography was visualized by ultraviolet light. Flash chromatography was performed on SilicaFlash$^{@}$

F60 silica gel (230–400 mesh). ^{1}H NMR spectra were recorded using a Varian Unity Inova 500 (500 MHz) or a Varian Unity Inova 300 (300 MHz). ^{13}C NMR spectra were recorded using a Varian Unity Inova 500 (125 MHz) or Varian Unity Inova 300 (75 MHz). The ^{1}H and ^{13}C NMR spectra were referenced to the residual solvent signals (7.26 ppm for ^{1}H and 77.0 ppm for ^{13}C in $CDCl_3$; 2.50 ppm for ^{1}H and 39.5 ppm for ^{13}C in DMSO-d6). All melting points were determined on a Büchi M560 melting point apparatus and uncorrected. Low (MS) and high (HRMS) mass spectra were determined on a Micromass Quattro II (ESI/APCI-TOF) at the University of Utah, Department of Chemistry Mass Spectrometry Facility.

2.4.14 Synthetic Routes, Procedures and Characterizations

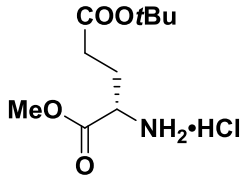


2.14

5-(*tert*-Butyl) 1-methyl (((9*H*-fluoren-9-yl)methoxy)carbonyl)-*L*-glutamate (2.14). To a stirred solution of 4.25 g Fmoc-Glu(O*t*Bu)-OH (10 mmol) in 40 mL DMF at room temperature, was added 1.40 g K_2CO_3 (10 mmol) and 0.75 mL CH_3I (12 mmol). The reaction was kept at room temperature for 4 h. Upon completion, the reaction mixture was poured into 100 mL 1 M HCl (aq) and extracted with ethyl acetate (3 ×50 mL). The organic layers were combined and washed successively by saturated $NaHCO_3$ (aq), 5% $Na_2S_2O_3$ (aq) and brine (100 mL each). The solvent was evaporated and *in vacuo* and the residue was purified by flash column chromatography (haxanes : ethyl acetate = 7:3) to yield 4.3 g product as colorless oil (98%). ^{1}H NMR (300 MHz, Chloroform-*d*) δ 7.77 (d, *J* = 7.5 Hz,

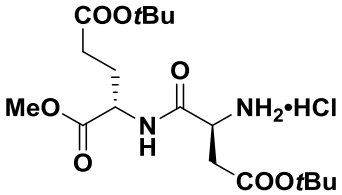
1H), 7.60 (dd, *J* = 7.6, 3.3 Hz, 2H), 7.40 (td, *J* = 7.4, 0.6 Hz, 1H), 7.32 (td, *J* = 7.7, 1.0 Hz, 1H), 5.50 (d, *J* = 8.2 Hz, 1H), 4.48 – 4.32 (m, 3H), 4.22 (t, *J* = 7.0 Hz, 1H), 3.76 (s, 3H), 2.33 (q, *J* = 7.4, 6.9 Hz, 2H), 2.28 – 2.07 (m, 1H), 1.98 (tt, *J* = 13.5, 6.2 Hz, 1H), 1.45 (s, 9H). ^{13}C NMR (75 MHz, $CDCl_3$) δ 172.8, 172.3, 156.2, 144.1, 143.9, 141.5, 127.9, 127.3, 125.3, 120.2, 81.1, 67.3, 53.7, 52.8, 47.4, 31.6, 28.3, 27.7.

General procedure for deprotection of Fmoc-protected amine. At 0 ˚C, to a stirred solution of Fmoc protected **53** in anhydrous dichloromethane (40 mL), equal volume of diethyl amine (40 mL) was added dropwise over 30 min. The reaction was kept at 0 ˚C for 4-6 h until TLC showed no starting material left. Upon completion, 20 mL anhydrous toluene was added to the reaction, and the mixture was evaporated under reduced pressure. The residue was purified by flash column chromatography (hexanes : ethyl acetate : methanol = 5:5:0 to 2:8:0.05) to afford the desired amine.



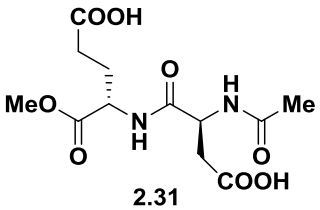
2.15

(*S*)-5-*tert*-Butyl 1-methyl 2-((*S*)-2-amino-4-(*tert*-butoxy)-4-oxobutanamido) pentanedioate (2.15). White solid, 65%. ^{1}H NMR (500 MHz, Chloroform-*d*) δ 3.65 (s, 3H), 3.40 (dd, *J* = 8.2, 5.3 Hz, 1H), 2.36 – 2.24 (m, 2H), 1.96 (dtd, *J* = 13.2, 7.6, 5.4 Hz, 1H), 1.78 – 1.66 (m, 1H), 1.37 (d, *J* = 1.6 Hz, 9H). ^{13}C NMR (126 MHz, *Chloroform-d*) δ 176.3, 172.5, 80.5, 53.9, 52.2, 31.9, 30.1, 28.2.



2.18

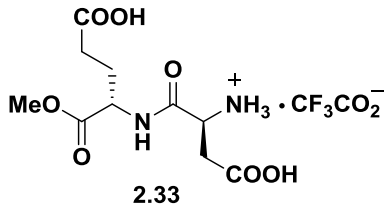
5-(*tert*-Butyl) 1-methyl ((S)-2-amino-4-(*tert*-butoxy)-4-oxobutanoyl)-*L*-glutamate hydrochloride (2.18). White solid, 85%. ^{1}H NMR (300 MHz, Chloroform-*d*) δ 7.90 (d, *J* = 8.4 Hz, 1H), 4.55 (td, *J* = 8.4, 5.0 Hz, 1H), 3.71 (s, 3H), 3.62 (dd, *J* = 7.8, 3.8 Hz, 1H), 2.75 (dd, *J* = 16.8, 3.8 Hz, 1H), 2.56 (dd, *J* = 16.8, 7.9 Hz, 1H), 2.27 (ddd, *J* = 7.6, 6.8, 3.6 Hz, 2H), 2.20 – 2.03 (m, 1H), 1.90 (dtd, *J* = 13.8, 8.3, 6.2 Hz, 1H), 1.41 (s, 9H), 1.40 (s, 9H). ^{13}C NMR (75 MHz, $CDCl_3$) δ 173.8, 172.5, 172.2, 171.4, 81.4, 80.8, 52.6, 52.2, 51.6, 40.6, 31.5, 28.3, 28.3, 27.8.



2.31

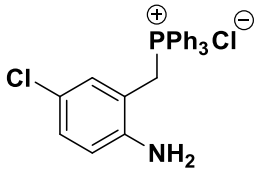
(*S*)-4-((*S*)-2-Acetamido-3-carboxypropanamido)-5-methoxy-5-oxopentanoic acid (2.31). To a stirred solution of **2.18** (0.219 g, 0.515 mmol) in 20 mL dichloromethane at 0 ˚C was added acetyl chloride (37 μL, 0.52 mmol) and triethyl amine (145 μL, 1.1 mmol). The reaction was kept at 0 ˚C for 3 h until TLC showed no starting material. The solvent was evaporated under reduced pressure. Resulting solids were broken up in 20 mL 1 M HCl (aq), and then extracted with ethyl acetate (2 × 15 mL). Organic layers were combined and washed with $NaHCO_3$ (aq) and brine successively, and then dried over Na_2SO_4. The solvent was evaporated under reduced pressure to afford 0.21 g **2.30** as white syrup (95%). Deprotection of the *tert*-Butyl esters followed the general procedure to yield

2.31 as white solid with quantitative yield. ^{1}H NMR (400 MHz, DMSO) δ ppm 8.77 (d, 1H, *J* = 4 Hz), 8.53 (d, 1H, *J* = 8 Hz), 4.59 (m. 1H), 4.26 (m, 1H), 3.62 (s, 3H), 2.62 (dd, 1H, *J* = 4.7 Hz, *J* = 16.5 Hz), 2.46 (dd, 1H, *J* = 8.6 Hz, *J* = 16.5 Hz), 2.28 (m, 2H), 1.96 (m, 1H), 1.80 (m, 4H). ^{13}C NMR (100 MHz, DMSO) δ ppm 174.2, 172.5, 172.0, 171.6, 169.8, 52.3, 51.8, 49.7, 36.7, 30.2, 26.4, 23.0. HRMS (ESI) *m/e* calcd. for $C_{12}H_{17}N_2O_8$ [M – H]$^-$ 317.0990, found 317.0987.



2.33

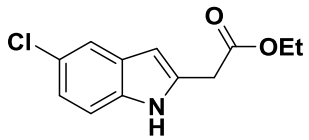
(*S*)-3-Carboxy-1-(((*S*)-4-carboxy-1-methoxy-1-oxobutan-2-yl) amino)-1-oxopropan-2-aminium 2,2,2-trifluoroacetate (2.33). White syrup; quantitative. ^{1}H NMR (500 MHz, DMSO) δ ppm 8.81 (d, 1H, *J* = 7 Hz), 8.19 (s, 3H), 4.33 (q, 1H, *J* = 8.2 Hz, *J* = 15.5 Hz), 4.11 (s, 1H), 3.62 (s, 3H), 2.81 (dd, 1H, *J* = 1.8 Hz, *J* = 18.3 Hz), 2.70 (dd, 1H, *J* = 8.8 Hz, *J* = 17.8 Hz), 2.29 (m, 2H), 1.99 (m, 1H), 1.79 (m, 1H). ^{13}C NMR (125 MHz, DMSO) δ ppm 174.2, 172.2, 171.5, 168.7, 52.8, 52.2, 49.4, 36.0, 30.3, 26.6. HRMS (ESI) *m/e* calcd. for $C_{10}H_{17}N_2O_7$ [M – CF_3COO]$^+$ 277.1030, found 277.1021.



2.23

(2-Amino-5-chlorobenzyl) triphenylphosphonium chloride (2.23). To a stirred solution of 3.15 g 5-chloro-2-amino benzyl alcohol (20 mmol), 5.25 g triphenylphosphine (20 mmol), and 5 g molecular sieves (4 Å) in 100 mL acetonitrile at reflux, was added 1.67 mL of 12 M HCl (aq) dropwise. A precipitate was formed, and the mixture was further

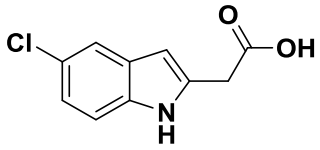
refluxed for 4 h until TLC showed no starting material. The reaction mixture was cooled in an ice bath allowing full precipitation. Solids were filtered. The filter cake was washed with ice cold acetonitrile (2 × 5 mL), hexanes (2 × 20 mL) to give 7.89 g product as light orange-pink solids (90%), which was pure enough to use for the next step without further purification.



2.24

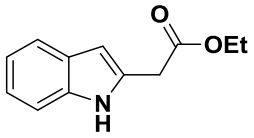
Ethyl 2-(5-chloro-1*H*-indol-2-yl) acetate (2.24). At room temperature, 4.4 g **2.23** (10 mmol) was suspended in 100 mL anhydrous dichloromethane. Then 1.54 mL ethyl 3-chloro-3-oxopropanoate (1.1 equiv) was added dropwise. The reaction was stirred at room temperature overnight. Upon completion, solvent was removed under reduced pressure to afford crude product, which was used without further purification. The generated (5-chloro-2-(3-ethoxy-3-oxopropanamido) benzyl) triphenyl phosphonium chloride (5.5 g, 10 mmol) was suspended in 150 mL anhydrous toluene under the protection of argon. The solution was heated to reflux for 15 min, and then 1.68 g *t*-BuOK (15 mmol) was added in small portions over 15 min. The reaction was further heated to reflux for another 30 min. Upon completion, the reaction was cooled to room temperature, and then filtered through a Celite plug. The resulting dark yellow-green filtrate was then concentrated under reduced pressure, and the residue was then purified by column chromatography (ethyl acetate : hexanes = 2:8) to afford 1.3 g product as a light orange solid (54% over 2 steps). ^{1}H NMR (300 MHz, *Chloroform-d*) δ ppm 8.84 (s, 1H), 7.51 (d, 1H, *J* = 2.0 Hz), 7.21 (td, 1H, *J* = 0.6 Hz, *J* = 8.6 Hz), 7.09 (dd, 1H, *J* = 2.0 Hz, *J* = 8.6 Hz), 6.29 (dd, 1H, *J* = 0.9 Hz, *J* =

2.0 Hz), 4.23 (q, 1H, *J* = 7.2 Hz), 3.81 (d, 1H, *J* = 0.7 Hz), 1.31 (t, 1H, *J* = 7.2 Hz). ^{13}C NMR (75 MHz, $CDCl_3$) δ ppm 170.8, 134.9, 132.4, 129.5, 125.6, 122.1, 119.7, 112.0, 101.7, 61.8, 34.0, 14.4. MS (ESI) *m*/*e* = 238.2 $[M + H]^+$.



2.25

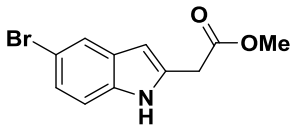
2-(5-Chloro-1*H*-indol-2-yl) acetic acid (2.25). To a stirred solution of 0.237 g **23a** (1 mmol) in 10 mL THF was added 10 mL 2 M LiOH (aq). The reaction was brought to reflux for 6 h. Upon completion, the reaction mixture was poured into 30 mL 1 M NaOH (aq), and then washed with dichloromethane (2 ×20 mL). The aqueous layer was acidified to pH ~4 by adding 2 M HCl (aq). A white precipitate was formed and extracted with ethyl acetate (2 ×30 mL). Organic layers were combined and washed with brine and dried over Na_2SO_4. Solvent was evaporated under reduced pressure to afford 0.201 g title compound as a light brown solid (96%). ^{1}H NMR (500 MHz) 12.54 (s, 1H), 11.20 (s, 1H), 7.47 (s, 1H), 7.31 (d, 1H, *J* = 8.6 Hz), 7.01 (m, 1H), 6.25 (d, 1H, *J* = 0.6 Hz), 3.74 (s, 1H). ^{13}C NMR (125 MHz) 172.0, 135.3, 135.2, 129.9, 123.9, 121.1, 119.2, 113.0, 101.1, 34.5. MS (ESI) *m*/*e* = 208.1 $[M - H]^+$.



2.43a

Ethyl 2-(1*H*-indol-2-yl) acetate (2.43a). The same procedure as **2.25**. Light brown oil; yield, 33% for four steps. ^{1}H NMR (300 MHz, *Chloroform-d*) δ ppm 8.75 (s, 1H), 7.64 (d, 1H, *J* = 7.5 Hz), 7.36 (d, 1H, *J* = 8.0 Hz), 7.24 (dd, 1H, *J* = 1.3 Hz, *J* = 7.1

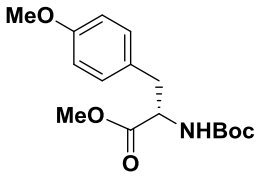
Hz), 7.20 (t, 1H, *J* = 1.8 Hz), 7.16 (dd, 1H, *J* = 1.2 Hz, *J* = 7.1 Hz), 6.42 (d, 1H, *J* = 1.1 Hz), 4.27 (q, 2H, *J* = 7.1 Hz), 3.85 (s, 2H), 1.35 (t, 3H, *J* = 7.2 Hz). ^{13}C NMR (75 MHz, $CDCl_3$) δ ppm 171.0, 136.6, 130.9, 128.5, 121.9, 120.4, 120.1, 111.2, 102.1, 61.7, 34.3, 14.5. MS (ESI) *m*/*e* = 204.1 $[M + H]^+$.



2.43b

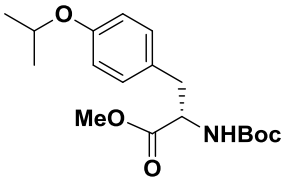
Methyl 2-(5-bromo-1*H*-indol-2-yl) acetate (2.43b). The same procedure as **2.25**. Light yellow oil; yield, 36% for four steps. ^{1}H NMR (300 MHz, *Chloroform-d*) δ ppm 8.75 (s, 1H), 7.50 (d, 1H, *J* = 1.9 Hz), 7.24 (d, 1H, *J* = 8.7 Hz), 7.10 (dd, 1H, *J* = 2.0 Hz, *J* = 8.6 Hz), 6.28 (dd, 1H, *J* = 0.9 Hz, *J* = 1.9 Hz), 3.83 (s, 2H), 3.76 (s, 3H). ^{13}C NMR (75 MHz, $CDCl_3$) δ ppm 171.1, 134.8, 132.1, 129.4, 125.6, 122.2, 119.7, 112.0, 101.8, 52.7, 33.8. MS (ESI) *m*/*e* = 268.1 $[M + H]^+$.

General procedure of alkylation of tyrosine. To a stirred solution of 0.295 g Boc-Tyr-OMe (1 mmol) in 15 mL acetone was added 0.15 g K_2CO_3 and 5 equiv of alkyl iodide. The reaction was heated to reflux overnight for MeI and 3 d for *i*-PrI. Upon completion, the reaction mixture was poured into 30 mL 1 M HCl (aq) and extracted with ethyl acetate (2 × 20 mL). Organic layers were combined and washed successively by saturated $NaHCO_3$ (aq), 5% $Na_2S_2O_3$ (aq), and brine. The organic layer was concentrated under reduced pressure, and the residue was purified by column chromatography to afford the desired alkylated product.



2.20

Boc-*L*-Tyr(OMe)-OMe (2.20). Colorless oil; yield, 92–95%. ^{1}H NMR (300 MHz, *Chloroform-d*) δ ppm 7.03 (d, 2H, *J* = 8.6 Hz), 6.82 (d, 2H, *J* = 8.7 Hz), 4.96 (d, 1H, *J* = 8.0 Hz), 4.53 (dd, 1H, *J* = 6.0 Hz, *J* = 14.2 Hz), 3.78 (s, 3H), 3.71 (s, 3H), 3.02 (m, 2H), 1.41 (s, 9H). ^{13}C NMR (75 MHz, $CDCl_3$) δ ppm 172.6, 158.8, 155.3, 130.5, 128.1, 114.2, 80.1, 55.4, 54.7, 52.4, 37.7, 28.5. MS (ESI) *m*/*e* = 310.2 $[M + H]^+$.

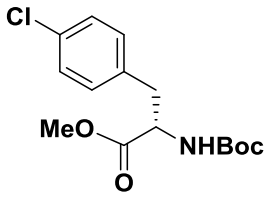


2.68

Boc-*L*-Tyr(OiPr)-OMe (2.68). Colorless oil; yield, 70–77%. ^{1}H NMR (500 MHz, *Chloroform-d*) δ ppm 6.99 (d, 2H, *J* = 8.2 Hz), 6.78 (d, 2H, *J* = 8.5 Hz), 4.99 (d, 1H, *J* = 7.6 Hz), 4.48 (m, 2H), 3.68 (s, 3H), 2.99 (ddd, 2H, *J* = 6.0 Hz, *J* = 14.0 Hz, *J* = 28.0 Hz), 1.39 (s, 9H), 1.29 (d, 6H, *J* = 6.1 Hz). ^{13}C NMR (125 MHz, $CDCl_3$) δ ppm 172.6, 157.1, 155.3, 130.5, 127.9, 116.1, 80.0, 70.0, 54.7, 52.3, 37.6, 28.5, 22.3. MS (ESI) *m*/*e* = 338.3 $[M + H]^+$.

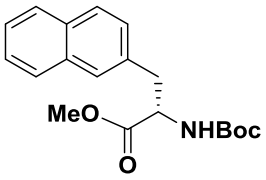
General procedure for formation of methyl ester. At 0 ˚C, to a stirred solution of carboxylic acid (1 equiv) and K_2CO_3 (1.2 equiv) in DMF was added MeI (1.5 equiv). The reaction was allowed to warm to room temperature. Upon completion, the reaction mixture was poured into 50 mL 2 M HCl (aq), and then extracted with ethyl acetate (2 × 30 mL). The organic layers were combined and washed with $NaHCO_3$ (aq), 5% $Na_2S_2O_3$

(aq), and brine. Solvent was removed under reduced pressure. Column chromatography was used to purify the product if needed.



2.74d

Boc-*L*-Phe(4-Cl)-OMe (2.74d). Colorless oil; yield, 97%. ^{1}H NMR (500 MHz, *Chloroform-d*) δ ppm 7.26 (d, 2H, *J* = 8.2 Hz), 7.05 (d, 2H, *J* = 8.1 Hz), 4.97 (d, 1H, *J* = 6.9 Hz), 4.57 (dd, 1H, *J* = 6.2 Hz, *J* = 13.2 Hz), 3.71 (s, 3H), 3.10 (dd, 1H, *J* = 5.6 Hz, *J* = 13.8 Hz), 3.00 (dd, 1H, *J* = 6.0 Hz, *J* = 13.9 Hz), 1.42 (s, 9H). ^{13}C NMR (125 MHz, $CDCl_3$) δ ppm 173.2, 172.3, 134.8, 130.8, 128.9, 95.0, 80.3, 54.5, 52.5, 38.0, 28.5.



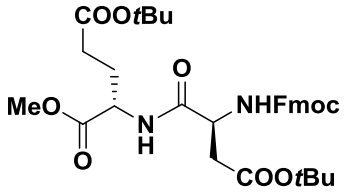
2.74e

Boc-*L*-(2-Nal)-OMe (2.74e). Pale yellow oil; yield, 92%. ^{1}H NMR (500 MHz, *Chloroform-d*) δ ppm 7.80 (m, 3H), 7.59 (s, 1H), 7.46 (m, 2H), 7.26 (m, 1H), 5.00 (d, 1H, *J* = 8.4 Hz), 4.67 (dd, 1H, *J* = 5.9 Hz, *J* = 13.6 Hz), 3.72 (s, 3H), 3.29 (dd, 1H, *J* = 5.9 Hz, *J* = 13.7 Hz), 3.21 (dd, 1H, *J* = 6.2 Hz, *J* = 13.9 Hz), 1.40 (s, 9H). ^{13}C NMR (125 MHz, $CDCl_3$) δ ppm 172.6, 155.3, 133.8, 133.6, 132.7, 128.5, 128.3, 127.9, 127.8, 127.5, 126.4, 125.9, 80.2, 54.7, 52.5, 38.7, 28.5.

General procedure for deprotection of Boc-protected amine. At 0 ˚C, to a solution of the Boc-protected amine in anhydrous dichloromethane (0.1 M, 20 mL) was added dropwise 20 mL 4 M HCl in dioxane. The reaction was kept at 0 ˚C for 10-12 h

until TLC showed no starting material left. Upon completion, the solvent was removed under reduced pressure to afford the desired amine as HCl salt. The amine was used directly without further purification.

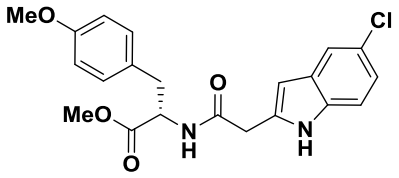
General peptide coupling procedure. At 0 ˚C, to a suspension of carboxylic acid (1 equiv), amine (HCl salt, 1 equiv), EDC•HCl (1.2 equiv), and HOAt (1 equiv) in anhydrous dichloromethane was added triethylamine (2.5 equiv) dropwise. A clear yellow solution was formed. The reaction was kept at 0 ˚C for 4-5 h until TLC showed no starting material left. The reaction was concentrated under reduced pressure. Residue was broken up in ethyl acetate and ice-cold 2 M HCl (aq). The organic layer was washed with equal volume of saturated $NaHCO_3$ (aq) and brine. The organic layer was dried over Na_2SO_4 and then evaporated under reduced pressure. Column chromatography was used to purify the final product if necessary.



2.17

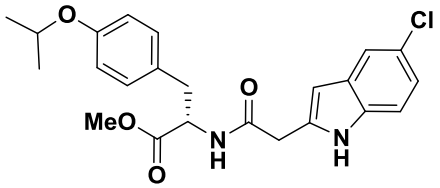
5-(*tert*-Butyl) 1-methyl ((*S*)-2-((((9H-fluoren-9-yl)methoxy)carbonyl)amino)-4-(*tert*-butoxy)-4-oxobutanoyl)-*L*-glutamate (2.17). ^{1}H NMR (300 MHz, Chloroform-d) δ 7.71 (d, J = 7.5 Hz, 1H), 7.56 (d, J = 7.4 Hz, 2H), 7.41 – 7.29 (m, 2H), 7.27 (t, J = 6.6 Hz, 2H), 6.07 (d, J = 8.7 Hz, 1H), 4.56 (td, J = 8.2, 4.8 Hz, 1H), 4.38 (d, J = 7.1 Hz, 2H), 4.19 (t, J = 7.1 Hz, 1H), 3.68 (s, 3H), 2.92 (dd, J = 17.3, 4.5 Hz, 1H), 2.60 (dd, J = 17.2, 6.3 Hz, 1H), 2.39 – 2.23 (m, 2H), 2.12 (d, J = 5.0 Hz, 0H), 1.93 (ddd, J = 12.7, 8.5, 6.2 Hz, 1H), 1.42 (s, 9H), 1.39 (s, 9H). ^{13}C NMR (75 MHz, $CDCl_3$) δ 172.3, 172.1, 170.8, 156.2, 143.8, 141.5, 127.9, 127.3, 125.3, 120.2, 82.1, 80.9, 77.8, 77.4, 77.0, 67.5, 52.7, 52.1, 51.3,

47.3, 37.8, 31.4, 28.2, 27.4.



2.26

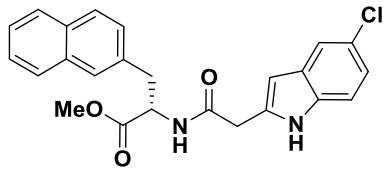
(*S*)-Methyl 2-(2-(5-chloro-1*H*-indol-2-yl) acetamido)-3-(4-methoxyphenyl) propanoate (2.26). Light orange solid; yield, 93%; m.p. 138.6-140.5˚C. ^{1}H NMR (300 MHz, *Chloroform-d*) δ ppm 9.07 (s, 1H), 7.51 (d, 1H, *J* = 2.0 Hz), 7.21 (d, 1H, *J* = 8.6 Hz), 7.11 (dd, 1H, *J* = 2.0 Hz, *J* = 8.6 Hz), 6.73 (d, 2H, *J* = 8.7 Hz), 6.54 (d, 2H, *J* = 8.7 Hz), 6.29 (t, 1H, *J* = 6.9 Hz), 6.23 (d, 1H, *J* = 1.3 Hz), 4.80 (td, 1H, *J* = 5.7 Hz, *J* = 7.9 Hz), 3.72 (s, 3H), 3.68 (s, 3H), 3.64 (s, 2H), 3.01 (dd, 1H, *J* = 5.7 Hz, *J* = 14.2 Hz), 2.94 (dd, 1H, *J* = 6.1 Hz, *J* = 14.2 Hz). ^{13}C NMR (75 MHz, $CDCl_3$) δ ppm 172.1, 169.4, 158.9, 135.0, 133.2, 130.2, 129.6, 127.2, 125.6, 122.3, 119.7, 114.1, 112.2, 101.8, 55.4, 53.5, 52.7, 36.8, 36.2. MS (ESI) *m*/*e* = 401.2 $[M + H]^+$.



2.75a

(*S*)-Methyl 2-(2-(5-chloro-1*H*-indol-2-yl) acetamido)-3-(4-isopropoxyphenyl) propanoate (2.75a). Light orange solid; yield, 87%. ^{1}H NMR (300 MHz, *Chloroform-d*) δ ppm 9.19 (s, 1H), 7.51 (d, 1H, *J* = 2.0 Hz), 7.21 (d, 1H, *J* = 8.6 Hz), 7.11 (dd, 1H, *J* = 2.0 Hz, *J* = 8.6 Hz), 6.70 (d, 2H, *J* = 8.6 Hz), 6.51 (d, 2H, *J* = 8.6 Hz), 6.36 (d, 1H, *J* = 7.8 Hz), 6.23 (d, 1H, *J* = 1.2 Hz), 4.79 (td, 1H, *J* = 5.7 Hz, *J* = 7.8 Hz), 4.36 (sept., 1H, *J* = 6.0 Hz), 3.71 (s, 3H), 3.64 (d, 2H, *J* = 1.5 Hz), 2.99 (dd, 1H, *J* = 5.3 Hz, *J* = 13.9 Hz),

2.93 (dd, 1H, *J* = 5.8 Hz, *J* = 14.0 Hz), 1.30 (d, 3H, *J* = 3.5 Hz), 1.28 (d, 3H, *J* = 3.6 Hz). ^{13}C NMR (75 MHz, $CDCl_3$) δ ppm 172.1, 169.5, 157.2, 135.0, 133.3, 130.3, 129.6, 126.9, 125.6, 122.3, 119.7, 116.0, 112.2, 101.8, 70.0, 53.5, 52.7, 36.8, 36.3, 22.3, 22.2.

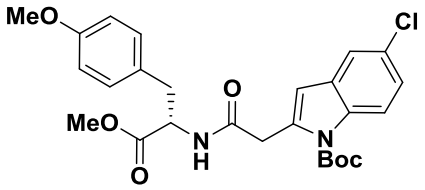


2.75e

(*S*)-Methyl 2-(2-(5-chloro-1*H*-indol-2-yl) acetamido)-3-(naphthalen-2-yl) propanoate (2.75e). Orange solid; yield, 91%. ^{1}H NMR (500 MHz, *Chloroform-d*) δ ppm 8.80 (s, 1H), 7.74 (d, 1H, *J* = 8.3 Hz), 7.56 (d, 1H, *J* = 8.4 Hz), 7.46 (m, 4H), 7.33 (s, 1H), 7.12 (s, 1H), 7.10 (s, 1H), 7.01 (d, 1H, *J* = 8.4 Hz), 6.26 (d, 1H, *J* = 7.6 Hz), 6.16 (s, 1H), 4.93 (dd, 1H, *J* = 6.1 Hz, *J* = 13.2 Hz), 3.74 (s, 3H), 3.60 (s, 2H), 3.27 (dd, 1H, *J* = 5.4 Hz, *J* = 13.9 Hz), 3.16 (dd, 1H, *J* = 6.3 Hz, *J* = 13.9 Hz). ^{13}C NMR (125 MHz, $CDCl_3$) δ ppm 172.0, 169.4, 134.9, 133.5, 132.9, 132.6, 129.5, 128.6, 128.2, 127.9, 127.7, 127.1, 126.5, 126.161, 125.7, 122.3, 119.8, 112.1, 101.7, 53.4, 52.8, 37.8, 36.2.

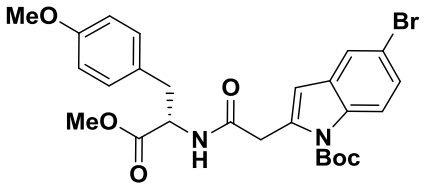
General procedure for formation of *N*-Boc protected indole. At room temperature, 1 equiv of indole compound was dissolved in anhydrous dichloromethane. Then 1.5 equiv of Boc_2O was added to the solution, followed by 0.1 equiv of DMAP. After 10-30 min, the reaction was confirmed to reach completion. The solvent was removed under reduced pressure, and the residue was purified by column chromatography (hexanes : ethyl acetate = 9:1). At 0 ˚C, to a solution of methyl ester in THF (0.1 M, 5 mL) was added dropwise 5 mL 2 M LiOH (aq). The reaction was kept at 0 ˚C for 2 h until TLC showed no starting material left. The reaction mixture was poured into 20 mL 2 M HCl (aq) and extracted with 20 mL ethyl acetate. The organic layer was then washed with 20

mL brine and dried over Na_2SO_4. Solvent was removed under reduced pressure to afford the desired carboxylic acid, which was used without further purification.



2.27

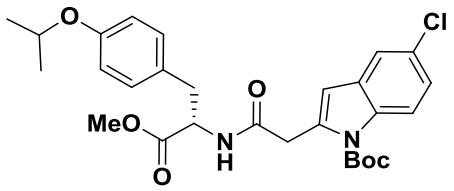
(*S*)-*tert*-Butyl 2-(2-((1-methoxy-3-(4-methoxyphenyl)-1-oxopropan-2-yl) amino)-2-oxoethyl)-1*H*-indole-1-carboxylate (2.27). Light yellow oil; yield, 86%. ^{1}H NMR (500 MHz, *Chloroform-d*) δ ppm 8.13 (d, 1H, *J* = 8.4 Hz), 7.50 (d, 1H, *J* = 7.8 Hz), 7.32 (t, 1H, *J* = 7.8 Hz), 7.24 (t, 1H, *J* = 8.1 Hz), 6.72 (d, 2H, *J* = 8.6 Hz), 6.53 (s, 1H), 6.46 (d, 2H, *J* = 8.6 Hz), 6.33 (d, 1H, *J* = 7.9 Hz), 4.82 (td, 1H, *J* = 5.5 Hz, *J* = 7.8 Hz), 3.97 (dd, 1H, *J* = 15.8 Hz, *J* = 36.3 Hz), 3.67 (s, 3H), 3.65 (s, 3H), 2.97 (dd, 1H, *J* = 5.6 Hz, *J* = 14.0 Hz), 2.93 (dd, 1H, *J* = 5.6 Hz, *J* = 14.0 Hz), 1.63 (s, 9H). ^{13}C NMR (125 MHz, $CDCl_3$) δ ppm 171.9, 169.1, 158.7, 150.8, 136.8, 134.0, 130.2, 129.1, 127.5, 124.5, 123.2, 120.7, 116.1, 114.0, 111.4, 85.2, 55.3, 53.2, 52.4, 38.3, 37.0, 28.3.



2.46b

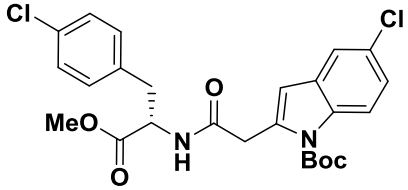
(*S*)-*tert*-Butyl 5-bromo-2-(2-((1-methoxy-3-(4-methoxyphenyl)-1-oxopropan-2-yl) amino)-2-oxoethyl)-1*H*-indole-1-carboxylate (2.46b). Light yellow oil; yield, 92%. ^{1}H NMR (500 MHz, *Chloroform-d*) δ ppm 7.98 (d, 1H, *J* = 8.9 Hz), 7.61 (d, 1H, *J* = 1.7 Hz), 7.40 (dd, 1H, *J* = 1.8 Hz, *J* = 8.9 Hz), 6.74 (d, 2H, *J* = 8.4 Hz), 6.49 (d, 2H, *J* = 8.4 Hz), 6.44 (s, 1H), 6.22 (d, 1H, *J* = 7.7 Hz), 4.81 (td, 1H, *J* = 5.5 Hz, *J* = 7.7 Hz),

3.94 (q, 2H, *J* = 15.9 Hz), 3.68 (s, 3H), 2.97 (dd, 1H, *J* = 5.8 Hz, *J* = 14.3 Hz), 2.93 (dd, 1H, *J* = 5.8 Hz, *J* = 14.2 Hz), 1.62 (s, 9H). ^{13}C NMR (125 MHz, $CDCl_3$) δ ppm 171.8, 168.7, 158.7, 150.3, 135.5, 135.3, 130.7, 130.2, 127.5, 127.2, 123.2, 117.5, 116.494, 114.1, 110.3, 85.7, 55.3, 53.2, 52.5, 38.2, 37.0, 28.3.



2.76a

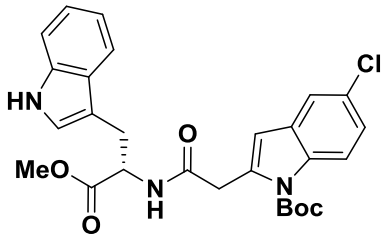
(*S*)-*tert*-Butyl 5-chloro-2-(2-((3-(4-isopropoxyphenyl)-1-methoxy-1-oxopropan -2-yl) amino)-2-oxoethyl)-1*H*-indole-1-carboxylate (2.76a). Light yellow oil; yield, 96%. ^{1}H NMR (300 MHz, *Chloroform-d*) δ ppm 8.04 (d, 1H, *J* = 8.9 Hz), 7.46 (d, 1H, *J* = 1.9 Hz), 7.26 (dd, 1H, *J* = 2.3 Hz, *J* = 8.7 Hz), 6.71 (d, 2H, *J* = 8.4 Hz), 6.46 (m, 3H), 6.19 (d, 1H, *J* = 8.1 Hz), 4.82 (td, 1H, *J* = 5.6 Hz, *J* = 8.2 Hz), 4.33 (sept, 1H, *J* = 6.0 Hz), 3.95 (dd, 2H, *J* = 15.9 Hz, *J* = 20.7 Hz), 3.68 (s, 3H), 2.98 (dd, 1H, *J* = 5.3 Hz, *J* = 13.9 Hz), 2.90 (dd, 1H, *J* = 5.6 Hz, *J* = 14.0 Hz), 1.63 (s, 9H), 1.28 (s, 3H), 1.26 (s, 3H). ^{13}C NMR (75 MHz, $CDCl_3$) δ ppm 171.9, 168.7, 157.1, 150.3, 135.4, 135.2, 130.2, 130.2, 128.8, 127.2, 124.5, 120.1, 117.1, 115.9, 115.9, 110.5, 85.7, 69.9, 53.1, 52.4, 38.3, 37.1, 28.3, 22.3, 22.2.



2.76d

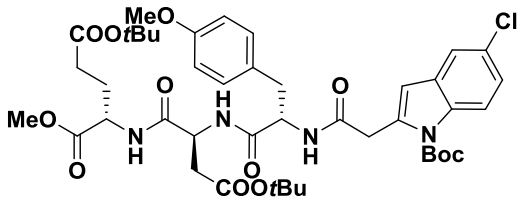
(*S*)-*tert*-Butyl 5-chloro-2-(2-((3-(4-chlorophenyl)-1-methoxy-1-oxopropan-2-yl) amino)-2-oxoethyl)-1*H*-indole-1-carboxylate (2.76d). Yellow oil; yield, 95%. ^{1}H

NMR (300 MHz, *Chloroform-d*) δ ppm 7.97 (d, 1H, *J* = 8.9 Hz), 7.46 (d, 1H, *J* = 2.0 Hz), 7.27 (dd, 1H, *J* = 2.2 Hz, *J* = 8.9 Hz), 6.94 (d, 2H, *J* = 8.4 Hz), 6.77 (d, 2H, *J* = 8.4 Hz), 6.50 (s, 1H), 6.48 (s, 1H), 4.81 (td, 1H, *J* = 5.6 Hz, *J* = 7.6 Hz), 3.95 (dd, 2H, *J* = 15.4 Hz, *J* = 52.7 Hz), 3.69 (s, 3H), 3.05 (dd, 1H, *J* = 5.6 Hz, *J* = 13.9 Hz), 2.97 (dd, 1H, *J* = 5.5 Hz, *J* = 13.9 Hz), 1.61 (s, 9H). ^{13}C NMR (75 MHz, $CDCl_3$) δ ppm 171.6, 168.9, 150.6, 135.6, 134.9, 134.3, 133.1, 130.6, 130.2, 128.9, 128.7, 124.7, 120.2, 117.2, 110.5, 85.7, 53.2, 52.6, 38.1, 37.1, 28.2.



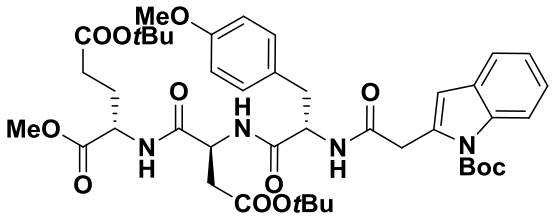
2.76f

(*S*)-*tert*-Butyl 2-(2-((3-(1*H*-indol-3-yl)-1-methoxy-1-oxopropan-2-yl) amino)-2-oxoethyl)-5-chloro-1*H*-indole-1-carboxylate (2.76f). Light yellow oil; yield, 93%. ^{1}H NMR (500 MHz, *Chloroform-d*) δ ppm 7.99 (d, 1H, *J* = 9.0 Hz), 7.88 (s, 1H), 7.43 (d, 1H, *J* = 8.0 Hz), 7.41 (d, 1H, *J* = 2.0 Hz), 7.26 (m, 2H), 7.13 (t, 1H, *J* = 7.6 Hz), 6.99 (dt, 1H, *J* = 0.9 Hz, *J* = 7.5 Hz), 6.67 (d, 1H, *J* = 1.8 Hz), 6.33 (s, 1H), 6.23 (d, 1H, *J* = 7.6 Hz), 4.87 (dd, 1H, *J* = 5.6 Hz, *J* = 13.2 Hz), 3.90 (q, 2H, *J* = 16.0 Hz), 3.65 (s, 3H), 3.23 (d, 2H, *J* = 5.6 Hz), 1.54 (s, 9H). ^{13}C NMR (125 MHz, $CDCl_3$) δ ppm 172.3, 168.9, 150.2, 136.2, 135.4, 135.2, 130.1, 128.7, 127.5, 124.5, 122.9, 122.4, 120.1, 119.8, 118.6, 117.0, 111.4, 110.4, 109.8, 85.5, 52.8, 52.6, 38.3, 28.2, 27.7.



2.29

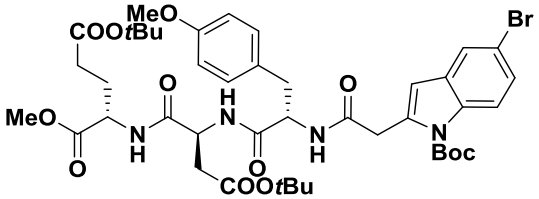
(*S*)-5-*tert*-Butyl 1-methyl 2-((*S*)-4-(*tert*-butoxy)-2-((S)-2-(2-(1-(*tert*-butoxycarbonyl)-5-chloro-1H-indol-2-yl) acetamido)-3-(4-methoxyphenyl) propanamido)-4-oxobutanamido) pentanedioate (2.29). Light yellow syrup; yield, 86%. ^{1}H NMR (300 MHz, *Chloroform-d*) δ ppm 8.01 (d, 1H, *J* = 9.0 Hz), 7.43 (d, 1H, *J* = 2.0 Hz), 7.28 (dd, 1H, *J* = 2.2 Hz, *J* = 9.0 Hz), 7.17 (dd, 1H, *J* = 3.1 Hz, *J* = 8.2 Hz), 6.76 (d, 2H, *J* = 8.6 Hz), 6.43 (d, 2H, *J* = 8.6 Hz), 6.13 (d, 1H, *J* = 6.4 Hz), 4.73 (ddd, 1H, *J* = 4.2 Hz, *J* = 6.1 Hz, *J* = 8.5 Hz), 4.50 (m, 2H), 3.92 (m, 2H), 3.71 (s, 3H), 3.66 (s, 3H), 2.93 (m, 3H), 2.46 (dd, 1H, *J* = 6.3 Hz, *J* = 17.2 Hz), 2.27 (m, 2H), 2.13 (m, 1H), 1.94 (m, 1H), 1.61 (s, 9H), 1.41 (s, 9H), 1.41 (s, 9H). ^{13}C NMR (75 MHz, $CDCl_3$) δ ppm 172.2, 172.0, 171.8, 170.5, 170.3, 170.0, 158.8, 150.2, 135.1, 134.8, 130.1, 130.0, 128.9, 127.4, 124.7, 120.1, 117.1, 114.2, 110.8, 85.8, 82.1, 80.8, 55.2, 55.0, 52.6, 52.1, 49.2, 38.2, 36.9, 36.4, 31.6, 28.3, 28.2, 27.3. MS (ESI) *m*/*e* = 857.4 $[M + H]^+$.



2.47a

(*S*)-5-*tert*-Butyl 1-methyl 2-((*S*)-4-(*tert*-butoxy)-2-((*S*)-2-(2-(1-(*tert*-butoxycarbonyl)-1*H*-indol-2-yl) acetamido)-3-(4-methoxyphenyl) propanamido)-4-oxobutanamido) pentanedioate (2.47a). Light orange syrup; yield, 83%. Two rotamers. ^{1}H NMR (500 MHz, *Chloroform-d*) δ ppm 8.08 (d, 0.6H, *J* = 8.3 Hz, major), 8.04 (d, 0.4H,

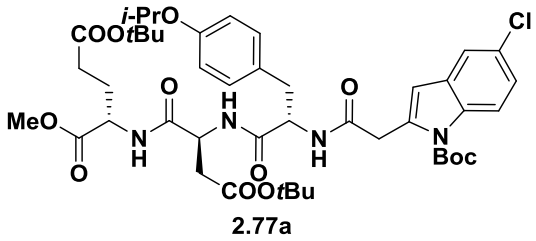
J = 8.3 Hz, minor), 7.47 (t, 1H, *J* = 8.6 Hz, major + minor), 7.27 (m, 0.8H, minor), 7.17 (dd, 1H, *J* = 5.1 Hz, *J* = 8.3 Hz, major + minor), 7.10 (d, 0.4H, *J* = 8.7 Hz, minor), 6.83 (d, 1H, *J* = 8.6 Hz, major), 6.72 (d, 1H, *J* = 8.6 Hz, major), 6.52 (m, 1H, minor), 6.48 (s, 0.6H, major), 6.38 (d, 1H, *J* = 8.6 Hz, minor), 6.34 (d, 0.4H, *J* = 6.3 Hz, minor), 6.17 (d, 0.6H, *J* = 6.3 Hz, major), 4.78 (m, 0.6H, major), 4.71 (m, 1H, major + minor), 4.49 (m, 2H, major + minor), 4.33 (q, 0.4H, *J* = 7.1 Hz, minor), 3.91 (m, 2H, major + minor), 3.70 (s, 1.8H, major), 3.69 (s, 1.2H, minor), 3.67 (s, 1.8H, major), 3.67 (s, 1.2H, minor), 3.62 (s, 2H), 2.87 (m, 3H, major + minor), 2.54 (dd, 0.6H, *J* = 6.8 Hz, *J* = 17.1 Hz, major), 2.47 (dd, 0.4H, *J* = 6.1 Hz, *J* = 17.2 Hz, minor), 2.32 (m, 1H, major + minor), 2.25 (m, 3H, major + minor), 2.13 (m, 1H), 1.93 (m, 1H), 1.60 (s, 9H), 1.40 (m, 18H, major + minor). ^{13}C NMR (125 MHz, $CDCl_3$) δ ppm 172.3, 172.1, 172.0, 172.0, 171.9, 171.8, 171.6, 170.9, 170.8, 170.5, 170.4, 170.4, 170.6, 170.3, 158.7, 158.7, 150.6, 136.7, 136.6, 133.6, 133.2, 130.1, 120.0, 129.0, 128.9, 127.9, 127.5, 124.7, 124.5, 123.3, 123.2, 120.6, 116.1, 116.0, 114.2, 111.8, 111.6, 85.3, 85.1, 82.1, 82.0, 81.0, 80.7, 56.0, 55.3, 55.2, 55.1, 52.6, 52.5, 52.1, 49.3, 49.2, 38.2, 37.3, 36.9, 36.6, 36.4, 36.3, 31.6, 31.5, 31.4, 28.3, 28.2, 27.3, 27.2, 23.4. MS (ESI) *m*/*e* = 823.4 $[M + H]^+$.



2.47b

(*S*)-5-*tert*-Butyl 1-methyl 2-((*S*)-2-((*S*)-2-(2-(5-bromo-1-(*tert*-butoxycarbonyl)-1H-indol-2-yl) acetamido)-3-(4-methoxyphenyl) propanamido)-4-(*tert*-butoxy)-4-oxobutanamido) pentanedioate (2.47b). Light yellow syrup; yield, 83%. ^{1}H NMR (500 MHz, *Chloroform-d*) δ ppm 7.98 (d, 1H, *J* = 9.0 Hz), 7.60 (d, 1H, *J* = 2.0 Hz), 7.43 (dd,

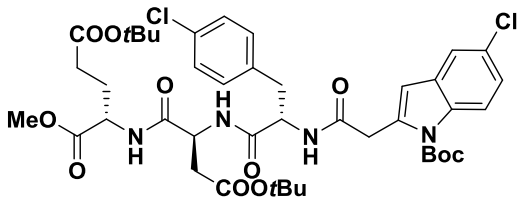
1H, *J* = 2.0 Hz, *J* = 8.9 Hz), 7.26 (m (overlapped with $CHCl_3$), 1H), 7.15 (d, 2H, *J* = 8.1 Hz), 6.77 (d, 2H, *J* = 8.6 Hz), 6.44 (d, 2H, *J* = 8.6 Hz), 6.43 (s, 1H), 6.10 (d, 1H, *J* = 6.5 Hz), 4.73 (ddd, 1H, *J* = 3.9 Hz, *J* = 6.0 Hz, *J* = 8.8 Hz), 4.51 (m, 2H), 3.93 (dd, 2H, *J* = 16.5 Hz, *J* = 28.5 Hz), 3.72 (s, 3H), 3.67 (s, 3H), 3.00 (dd, 1H, *J* = 6.7 Hz, *J* = 14.1 Hz), 2.92 (m, 3H), 2.46 (dd, 1H, *J* = 6.3 Hz, *J* = 17.2 Hz), 2.28 (m, 2H), 2.15 (m, 1H), 1.95 (m, 1H), 1.62 (s, 9H), 1.42 (s, 9H), 1.42 (s, 9H). ^{13}C NMR (125 MHz, $CDCl_3$) δ ppm 172.1, 171.9, 171.8, 170.5, 170.3, 170.1, 169.9, 158.8, 135.5, 134.7, 130.6, 130.4, 130.0, 127.5, 127.4, 123.2, 117.5, 116.6, 114.3, 110.7, 85.9, 82.1, 55.3, 55.0, 52.6, 52.1, 49.2, 38.2, 36.9, 36.5, 31.6, 28.3, 28.3, 28.2, 27.4. MS (ESI) *m*/*e* = 901.1 $[M + H]^+$.



2.77a

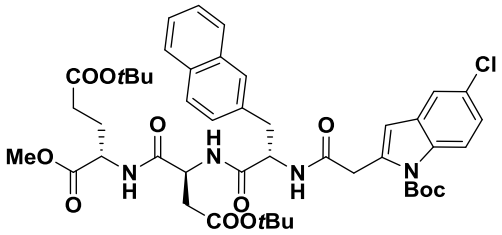
(*S*)-5-*tert*-Butyl 1-methyl 2-((*S*)-4-(*tert*-butoxy)-2-((*S*)-2-(2-(1-(*tert*-butoxycarbonyl)-5-chloro-1*H*-indol-2-yl) acetamido)-3-(4-isopropoxyphenyl) propanamido)-4-oxobutanamido) pentanedioate (2.77a). Light orange oil; yield, 76%. Two rotamers. ^1H NMR (500 MHz, *Chloroform-d*) δ ppm 8.03 (d, 0.6H, *J* = 9.0 Hz, major), 7.99 (d, 0.4H, *J* = 9.0 Hz, minor), 7.47 (d, 0.4H, *J* = 2.1 Hz, minor), 7.44 (d, 0.6H, *J* = 2.0 Hz, major), 7.30 (m, 1H, major + minor), 7.15 (dd, 1.2H, *J* = 5.5 Hz, *J* = 8.2 Hz, major), 7.06 (dd, 0.8H, *J* = 8.6 Hz, *J* = 18.8 Hz, minor), 6.87 (d, 0.8H, *J* = 8.5 Hz, minor), 6.75 (d, 1.2H, *J* = 8.5 Hz, major), 6.56 (d, 0.8H, *J* = 8.6 Hz, minor), 6.48 (s, 0.4H, mimor), 6.43 (d, 1.2H, *J* = 5.6 Hz, major), 6.40 (s, 0.8H, major), 6.28 (d, 0.4H, *J* = 6.6 Hz, minor), 6.12 (d, 0.6H, *J* = 6.4 Hz, major), 4.71 (m, 1H, major + minor), 4.50 (m, 2H, major + minor), 4.40 (m, 1H, major + minor), 4.33 (sept, 1H, *J* = 6.0 Hz), 3.93 (m, 2H, major +

minor), 3.72 (s, 2H, major), 3.69 (s, 1H, minor), 3.01 (dd, 1H, *J* = 6.5 Hz, *J* = 14.0 Hz), 2.90 (m, 1H), 2.45 (dd, 1H, *J* = 6.3 Hz, *J* = 17.2 Hz), 2.29 (m, 3H), 2.13 (m, 1H), 1.94 (m, 1H), 1.63 (s, 9H), 1.42 (s, 9H), 1.42 (s, 9H). ^{13}C NMR (125 MHz, $CDCl_3$) δ ppm 172.1, 171.9, 171.8, 170.3, 170.0, 150.2, 130.2, 130.1, 130.0, 127.2, 124.7, 124.6, 120.2, 117.1, 117.1, 116.2, 116.1, 110.9, 110.7, 85.8, 85.6, 82.1, 80.8, 69.8, 55.9, 54.9, 52.6, 52.2, 52.1, 49.2, 38.2, 36.9, 36.5, 31.6, 28.3, 28.3, 28.2, 28.2, 28.2, 27.3, 27.2, 22.3, 22.2, 22.1. MS (ESI) *m*/*e* = 885.3 $[M + H]^+$.



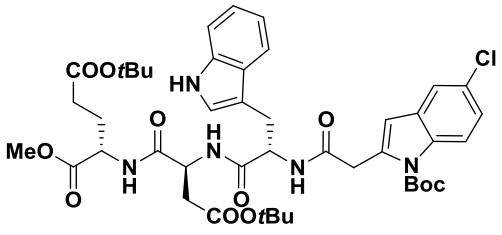
2.77d

(*S*)-5-*tert*-Butyl 1-methyl 2-((*S*)-4-(*tert*-butoxy)-2-((*S*)-2-(2-(1-(*tert*-butoxycarbonyl)-5-chloro-1*H*-indol-2-yl) acetamido)-3-(4-chlorophenyl) propanamido)-4-oxobutanamido) pentanedioate (2.77d). Orange syrup; yield, 82%. ^{1}H NMR (500 MHz, *Chloroform-d*) δ ppm 7.95 (d, 1H, *J* = 8.8 Hz), 7.46 (d, 1H, *J* = 1.6 Hz), 7.28 (m, 2H), 7.19 (m, 2H), 6.92 (d, 2H, *J* = 8.1 Hz), 6.85 (d, 2H, *J* = 8.2 Hz), 6.45 (s, 1H), 6.29 (d, 1H, *J* = 6.6 Hz), 4.71 (td, 1H, *J* = 6.9 Hz, *J* = 13.5 Hz), 4.52 (m, 2H), 3.91 (s, 2H), 3.71 (s, 3H), 2.99 (m, 2H), 2.89 (dd, 1H, *J* = 3.9 Hz, *J* = 17.2 Hz), 2.51 (dd, 1H, *J* = 6.4 Hz, *J* = 17.2 Hz), 2.29 (m, 2H), 2.13 (m, 1H), 1.93 (m, 1H), 1.61 (s, 1H), 1.43 (s, 9H), 1.42 (s, 9H). ^{13}C NMR (125 MHz, $CDCl_3$) δ ppm 172.2, 171.9, 171.7, 170.3, 170.2, 169.9, 150.3, 134.9, 134.9, 134.5, 133.2, 130.5, 130.4, 130.0, 129.0, 129.0, 124.8, 120.2, 117.1, 110.7, 85.8, 82.2, 80.9, 54.8, 52.6, 52.1, 49.4, 38.0, 37.1, 36.8, 31.6, 28.3, 28.2, 27.3. MS (ESI) *m*/*e* = 861.5 $[M + H]^+$.



2.77e

(*S*)-5-*tert*-Butyl 1-methyl 2-((*S*)-4-(*tert*-butoxy)-2-((*S*)-2-(2-(5-chloro-1*H*-indol-2-yl) acetamido)-3-(naphthalen-2-yl) propanamido)-4-oxobutanamido) pentanedioate (2.77e). Orange syrup; yield, 86%. ^{1}H NMR (500 MHz, *Chloroform-d*) δ ppm 9.52 (s, 0.5H), 9.31 (s, 0.5H), 7.74 (m, 1H), 7.64 (m, 1H), 7.46 (m, 6H), 7.33 (m, 2H), 7.20 (m, 1H), 7.06 (m, 2H), 6.67 (d, 0.5H, *J* = 6.3 Hz), 6.44 (d, 0.5H, *J* = 6.3 Hz), 6.12 (s, 1H), 4.73 (m, 2H), 4.53 (m, 2H), 3.72 (s, 0.3H), 3.69 (s, 1.7H), 3.66 (m, 3H), 3.16 (m, 2H), 2.87 (m, 1H), 2.42 (dd, 0.5H, *J* = 6.3 Hz, *J* = 17.2 Hz), 2.26 (m, 2.5H), 2.12 (m, 2H), 1.92 (m, 1H), 1.41 (s, 9H), 1.34 (s, 9H). ^{13}C NMR (125 MHz, $CDCl_3$) δ ppm 172.7, 172.6, 172.5, 172.0, 171.6, 171.0, 170.7, 170.5, 170.4, 135.3, 135.2, 133.6, 133.5, 133.1, 132.9, 132.8, 132.6, 129.5, 129.4, 128.7, 128.7, 128.1, 128.0, 127.9, 127.7, 127.1, 126.9, 126.6, 126.2, 125.7, 125.4, 122.4, 122.1, 119.8, 119.6, 112.4, 112.3, 102.1, 101.8, 82.3, 82.1, 81.2, 81.0, 56.1, 55.2, 52.7, 52.1, 52.0, 49.4, 49.3, 37.5, 37.3, 37.0, 36.4, 36.3, 36.0, 31.8, 31.6, 28.3, 28.2, 27.6, 27.5. MS (ESI) *m*/*e* = 777.2 $[M + H]^+$.

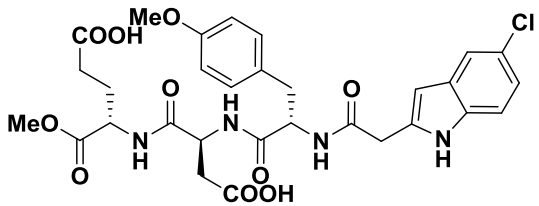


2.77f

(*S*)-5-*tert*-Butyl 1-methyl 2-((*S*)-4-(*tert*-butoxy)-2-((*S*)-2-(2-(1-(*tert*-butoxycarbonyl)-5-chloro-1*H*-indol-2-yl) acetamido)-3-(1*H*-indol-3-yl)

propanamido)-4-oxobutanamido) pentanedioate (2.77f). Light orange syrup; yield, 78%. ^{1}H NMR (500 MHz, *Chloroform-d*) δ ppm 8.14 (s, 1H), 7.99 (d, 1H, *J* = 8.9 Hz), 7.59 (d, 1H, *J* = 7.9 Hz), 7.44 (d, 1H, *J* = 2.1 Hz), 7.28 (dd, 1H, *J* = 2.1 Hz, *J* = 8.9 Hz), 7.25 (d, 1H, *J* = 8.2 Hz), 7.13 (t, 1H, *J* = 7.6 Hz), 6.99 (m, 2H), 6.84 (d, 1H, *J* = 8.4 Hz), 6.76 (d, 1H, *J* = 2.3 Hz), 6.42 (s, 1H), 6.36 (d, 1H, *J* = 6.4 Hz), 4.67 (m, 2H), 4.48 (dt, 1H, *J* = 5.3 Hz, *J* = 8.3 Hz), 3.94 (m, 2H), 3.71 (s, 3H), 3.28 (dd, 1H, *J* = 5.6 Hz, *J* = 14.5 Hz), 3.11 (dd, 1H, *J* = 7.7 Hz, *J* = 14.5 Hz), 2.82 (dd, 1H, *J* = 4.0 Hz, *J* = 17.2 Hz), 2.33 (dd, 1H, *J* = 6.4 Hz, *J* = 17.2 Hz), 2.25 (m, 2H), 2.09 (m, 1H), 1.88 (m, 1H), 1.57 (s, 9H), 1.44 (s, 9H), 1.36 (s, 9H). ^{13}C NMR (125 MHz, $CDCl_3$) δ ppm 172.5, 172.0, 171.5, 171.1, 170.2, 169.8, 150.2, 136.5, 135.2, 135.1, 130.2, 128.8, 126.9, 124.6, 123.1, 122.7, 120.1, 120.0, 118.9, 117.1, 111.5, 110.7, 110.1, 85.6, 82.0, 81.1, 54.2, 52.7, 51.9, 49.3, 38.3, 36.8, 31.5, 29.9, 28.3, 28.3, 28.2, 27.5. MS (ESI) *m*/*e* = 866.5 $[M + H]^+$.

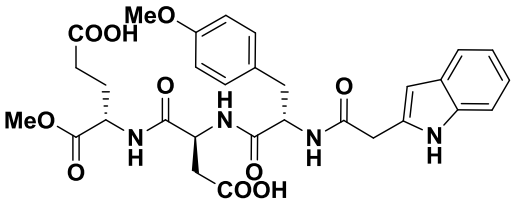
General procedure for deprotection of *tert*-Butyl ester. At 0 ˚C degree, to a solution of *tert*-butyl ester in anhydrous dichloromethane (0.1 M 10mL) was added 10 mL of trifluoroacetic acid dropwise. The reaction was kept at 0 ˚C for 4–6 h until TLC showed no starting material left. Upon completion, 10 mL of anhydrous toluene was added, and the solvent was removed under reduced pressure to afford desired product as light pink solid.



2.12

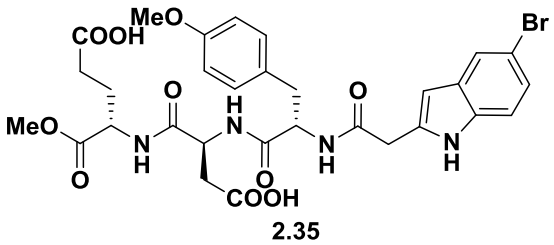
(*S*)-4-((*S*)-3-Carboxy-2-((*S*)-2-(2-(5-chloro-1*H*-indol-2-yl) acetamido)-3-(4-methoxyphenyl) propanamido) propanamido)-5-methoxy-5-oxopentanoic acid (2.12).

Light pink solid; yield, 93%; m.p., decompose at 118 ˚C. ^{1}H NMR (500 MHz, DMSO) δ ppm 12.24 (br, 2H), 11.02 (s, 1H), 8.45 (t, 1H, *J* = 7.3 Hz), 8.19 (m, 2H), 7.41 (s, 1H), 7.28 (d, 1H, *J* = 8.6 Hz), 7.04 (t, 2H, *J* = 7.9 Hz), 6.98 (dd, 1H, *J* = 1.5 Hz, *J* = 8.9 Hz), 6.65 (dd, 2H, *J* = 2.4 Hz, *J* = 8.4 Hz), 6.00 (d, 1H, *J* = 18.6 Hz), 4.62 (m, 1H), 4.56 (m, 0.5H), 4.45 (m, 1H), 4.24 (m, 1.5H), 3.63 (d, 3H, *J* = 2.9 Hz), 3.59 (m, 5H), 3.50 (m, 1H), 2.90 (m, 1H), 2.64 (m, 2H), 2.41 (dd, 1H, *J* = 8.2 Hz, *J* = 16.3 Hz), 2.26 (m, 3H), 1.92 (m, 1H), 1.78 (m, 1H). ^{13}C NMR (500 MHz, DMSO) 174.4, 172.6, 172.2, 171.7, 171.5, 169.2, 158.4, 136.3, 135.1, 130.9, 129.9, 139.6, 128.9, 123.9, 120.8, 119.1, 114.0, 113.0, 100.5, 100.4, 55.5, 55.1, 52.6, 52.4, 52.0, 49.8, 37.5, 37.4, 36.9, 36.7, 35.9, 30.4, 26.7, 26.5. HRMS (ESI) *m/e* calcd. for $C_{30}H_{32}ClN_4O_{10}$ $[M - H]^-$ 643.1812, found 643.1819.



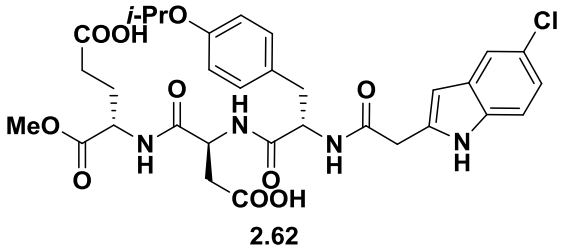
2.34

(*S*)-4-((*S*)-2-((*S*)-2-(2-(1*H*-Indol-2-yl) acetamido)-3-(4-methoxyphenyl) propanamido)-3-carboxypropanamido)-5-methoxy-5-oxopentanoic acid (2.34). Light purple solid; yield, quantitative; m.p., decompose at 78 ℃. ^{1}H NMR (500 MHz, DMSO) δ ppm 12.25 (br, 2H), 10.79 (s, 1H), 8.44 (m, 1H), 8.16 (m, 2H), 7.37–6.92 (m, 6H), 6.65 (d, 2H, J = 8 Hz), 6.03 (s, 1H), 4.57–4.47 (m, 2H), 4.27 (m, 1H), 3.70–3.52 (m, 8H), 2.92 (m, 1H), 2.67 (d, 1H, J = 12 Hz), 2.29 (m, 2H), 1.94 (m, 1H), 1.80 (m, 1H). ^{13}C NMR (125 MHz, DMSO) δ ppm 174.4, 172.7, 172.3, 172.0, 171.4, 169.3, 158.4, 136.7, 134.3, 130.9, 130.1, 129.6, 128.9, 128.8, 126.0, 121.0, 119.9, 119.3, 114.0, 111.5, 100.6, 55.5, 54.8, 52.6, 52.0, 50.1, 37.5, 36.7, 36.0, 30.4, 26.7. HRMS (ESI) *m/e* calcd. for $C_{30}H_{33}N_4O_{10}$ $[M - H]^-$ 609.2202, found 609.2217.



2.35

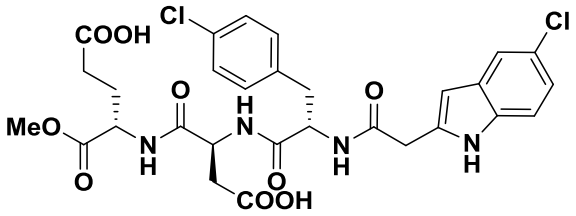
(*S*)-4-((*S*)-2-((*S*)-2-(2-(5-Bromo-1*H*-indol-2-yl) acetamido)-3-(4-methoxyphenyl) propanamido)-3-carboxypropanamido)-5-methoxy-5-oxopentanoic acid (2.35). Light pink solid; yield, 97%; m.p., decompose at 166 ˚C. ^{1}H NMR (500 MHz, DMSO) δ ppm 11.02 (s, 1H), 8.44 (d, 1H, *J* = 6.6 Hz), 8.19 (d, 1H, *J* = 7.7 Hz), 8.15 (d, 1H, *J* = 6.9 Hz), 7.54 (s, 1H), 7.24 (d, 1H, *J* = 7.9 Hz), 7.09 (d, 1H, *J* = 8.6 Hz), 7.04 (d, 2H, *J* = 7.8 Hz), 6.64 (d, 2H, *J* = 7.9 Hz), 5.97 (s, 1H), 4.55 (m, 1H), 4.44 (dd, 1H, *J* = 1.3 Hz, *J* = 5.9 Hz), 4.26 (m, 1H), 3.62 (s, 3H), 3.59 (s, 3H), 2.93 (d, 2H, *J* = 14.6 Hz), 2.64 (m, 1H), 2.25 (m, 1H), 1.91 (m, 1H), 1.76 (m, 1H). ^{13}C NMR (125 MHz, DMSO) δ ppm 174.4, 172.7, 172.2, 171.9, 136.1, 135.3, 130.9, 130.7, 130.1, 123.4, 122.1, 114.1, 114.0, 113.5, 111.9, 100.3, 95.0, 55.4, 54.8, 54.7, 52.6, 52.0, 50.1, 37.5, 36.7, 35.9, 30.4, 26.7. HRMS (ESI) *m/e* calcd. for $C_{30}H_{32}BrN_4O_{10}$ $[M - H]^-$ 687.1307, found 687.1309.



2.62

(*S*)-4-((*S*)-3-carboxy-2-((*S*)-2-(2-(5-chloro-1*H*-indol-2-yl)acetamido)-3-(4-isopropoxyphenyl)propanamido)propanamido)-5-methoxy-5-oxopentanoic acid (2.62). Pink solid; yield, quantitative; m.p., decompose at 106 ˚C. ^{1}H NMR (500 MHz, DMSO) δ ppm 12.24 (br, 2H), 11.05 (d, 1H, *J* = 9.6 Hz), 8.45 (t, 1H, *J* = 7.1 Hz), 8.26 (dd, 1H, *J* = 8.1 Hz, *J* = 23.6 Hz), 8.18 (dd, 1H, *J* = 7.5 Hz, *J* = 18.9 Hz), 7.41 (s, 1H),

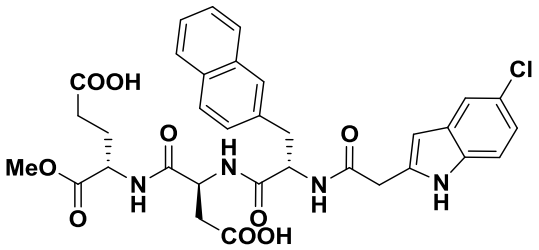
7.28 (d, 1H, *J* = 8.5 Hz), 7.23 (t, 1H, *J* = 7.5 Hz), 7.14 (dd, 1H, *J* = 7.2 Hz, *J* = 18.1 Hz), 7.01 (m, 2H), 6.97 (dd, 1H, *J* = 1.3 Hz, *J* = 8.6 Hz), 6.60 (d, 2H, *J* = 7.7 Hz), 5.98 (d, 1H, *J* = 20.5 Hz), 4.62 (dt, 0.5H, *J* = 5.2 Hz, *J* = 8.3 Hz), 4.56 (dt, 0.5H, *J* = 5.3 Hz, *J* = 8.0 Hz), 4.42 (m, 2H) 4.24 (m, 1H), 3.60 (s, 2H), 3.58 (s, 3H), 3.49 (dd, 2H, *J* = 11.4 Hz, *J* = 15.1 Hz), 2.89 (ddd, 1H, *J* = 4.1 Hz, *J* = 13.8 Hz, *J* = 24.5 Hz), 2.64 (m, 2H), 2.40 (dd, 1H, *J* = 8.5 Hz, *J* = 16.5 Hz), 2.26 (m, 3H), 1.91 (m, 1H), 1.77 (m, 1H), 1.20 (s, 3H), 1.19 (s, 3H). ^{13}C NMR (125 MHz, DMSO) δ ppm 174.5, 174.4, 172.7, 172.6, 172.3, 172.2, 170.0, 171.7, 171.5, 171.4, 169.2, 169.0, 156.6, 156.5, 136.3, 136.3, 135.1, 130.9, 130.9, 130.0, 129.9, 129.9, 129.7, 129.6, 128.9, 126.0, 123.9, 120.8, 119.1, 115.6, 115.6, 112.9, 100.4, 100.4, 69.5, 69.5, 55.1, 54.8, 52.6, 52.6, 52.2, 52.0, 50.1, 49.8, 35.9, 30.5, 30.4, 26.7, 26.5, 22.6, 22.5, 21.7. HRMS (ESI) *m/e* calcd. for $C_{32}H_{36}ClN_4O_{10}$ $[M - H]^-$ 671.2125, found 671.2129.



2.65

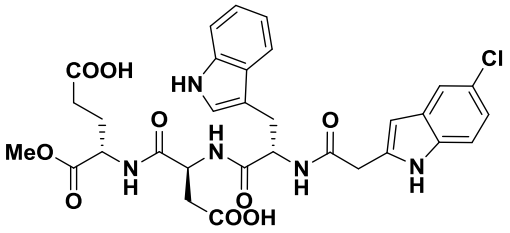
(*S*)-4-((*S*)-3-Carboxy-2-((*S*)-2-(2-(5-chloro-1*H*-indol-2-yl) acetamido)-3-(4-chlorophenyl) propanamido) propanamido)-5-methoxy-5-oxopentanoic acid (2.65). Light orange solid; yield, 98%; m.p., decompose at 196 ˚C. ^{1}H NMR (500 MHz, DMSO) δ ppm 12.25 (br, 2H), 11.02 (s, 1H) 8.48 (d, 1H, *J* = 7.6 Hz), 8.23 (dd, 2H, *J* = 8.0 Hz, *J* = 24.1 Hz), 7.42 (d, 1H, *J* = 1.4 Hz), 7.28 (d, 1H, *J* = 8.6 Hz), 7.15 (m, 4H), 6.98 (dd, 1H, *J* = 2.1 Hz, *J* = 8.6 Hz), 5.97 (s, 1H), 4.55 (m, 2H), 4.27 (m, 1H), 3.60 (s, 3H), 3.57 (d, 2H, *J* = 14.0 Hz), 3.49 (d, 1H, *J* = 15.2 Hz), 2.99 (dd, 1H, *J* = 3.9 Hz, *J* = 13.9 Hz), 2.68

(m, 3H), 2.51 (d, 1H, *J* = 8.9 Hz), 2.27 (m, 2H), 1.93 (m, 1H), 1.77 (m, 1H). ^{13}C NMR (125 MHz, DMSO) δ ppm 174.4, 172.7, 172.2, 171.6, 171.4, 169.0, 137.3, 136.3, 135.1, 131.8, 131.6, 129.9, 128.5, 123.9, 120.8, 119.1, 112.9, 100.4, 95.0, 54.2, 52.6, 52.0, 50.2, 36.8, 35.9, 30.4, 26.7. HRMS (ESI) *m/e* calcd. for $C_{29}H_{29}Cl_2N_4O_9$ $[M - H]^-$ 647.1317, found 647.1317.



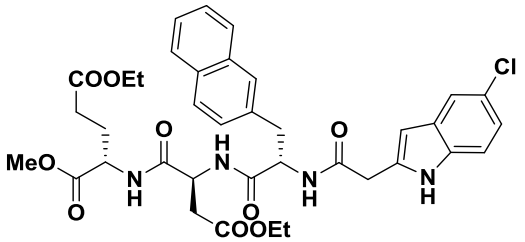
2.66

(*S*)-4-((*S*)-3-Carboxy-2-((*S*)-2-(2-(5-chloro-1H-indol-2-yl) acetamido)-3-(naphthalen-2-yl) propanamido) propanamido)-5-methoxy-5-oxopentanoic acid (2.66). Light pink solid; yield, quantitative; m.p., decompose at 137 ˚C. ^{1}H NMR (400 MHz, DMSO) δ ppm 12.26 (br, 2H), 10.73 (s, 1H), 8.25(m, 3H), 6.88–7.86(m, 11H), 6.03(s, 1H) 4.67(m, 2H), 4.32(m, 1H), 3.64(m, 5H), 2.95(m, 2H), 2.68(m, 2H), 2.30(m, 4H), 1.94(m, 1H), 1.81(m, 1H). ^{13}C NMR (125 MHz, DMSO) δ ppm 174.4, 172.7, 172.6, 172.4, 172.0, 171.6, 171.0, 170.7, 170.5, 170.4, 135.3, 135.2, 133.6, 133.5, 133.1, 132.9, 132.8, 132.6, 129.5, 129.4, 128.7, 128.7, 128.1, 128.0, 127.9, 127.7, 127.1, 126.9, 126.6, 126.2, 125.7, 125.4, 122.4, 122.1, 119.8, 119.6, 112.4, 112.3, 102.1, 101.8, 56.1, 55.2, 52.7, 52.1, 52.0, 49.4, 49.3, 37.5, 37.3, 37.0, 36.4, 36.3, 36.0, 31.8, 31.6, 26.5. HRMS (ESI) *m/e* calcd. for $C_{33}H_{32}ClN_4O_9$ $[M - H]^-$ 663.1863, found 663.1869.



2.67

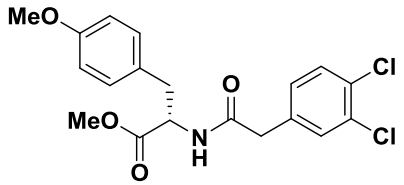
(*S*)-4-((*S*)-3-Carboxy-2-((*S*)-2-(2-(5-chloro-1H-indol-2-yl) acetamido)-3-(1H-indol-3-yl) propanamido) propanamido)-5-methoxy-5-oxopentanoic acid (2.67). Light pink solid; yield, 96%; m.p., decompose at 128 ˚C. ^{1}H NMR (500 MHz, DMSO) δ ppm 12.24 (br, 1H), 11.00 (m, 1H), 10.79 (s, 1H), 8.36 (m, 3H), 7.59 (dd, 1H, *J* = 7.2 Hz, *J* = 13.5 Hz), 7.39 (d, 1H, *J* = 11.0 Hz), 7.27 (m, 4H), 7.15 (m, 3H), 7.04 (t, 1H, *J* = 7.2 Hz), 6.94 (m, 2H), 5.96 (m, 1H), 4.58 (m, 2H), 4.24 (m, 1H), 3.58 (m, 5H), 3.11 (dd, 1H, *J* = 10.0 Hz, *J* = 24.6 Hz), 2.92 (m, 1H), 2.64 (m, 2H), 2.30 (m, 2H), 1.91 (m, 1H), 1.77 (m, 1H). ^{13}C NMR (125 MHz, DMSO) δ ppm 174.4, 172.7, 172.6, 172.3, 172.2, 171.5, 171.4, 169.2, 140.7, 138.0, 136.7, 136.7, 136.3, 135.1, 129.9, 129.6, 128.9, 128.0, 126.0, 124.4, 123.8, 121.5, 120.8, 119.2, 119.1, 119.0, 118.8, 112.9, 112.9, 111.9, 110.6, 110.4, 100.4, 54.2, 52.6, 52.5, 52.1, 52.0, 50.2, 49.8, 36.6, 35.7, 35.6, 30.4, 30.3, 28.5, 26.7, 26.5. HRMS (ESI) *m/e* calcd. for $C_{31}H_3ClN_5NaO_9$ $[M - 2H + Na]^-$ 674.1635, found 674.1654.



2.78

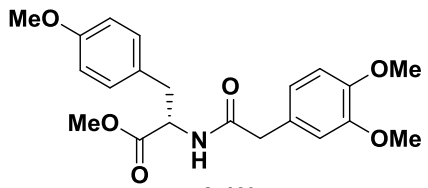
(*S*)-5-Ethyl 1-methyl 2-((*S*)-2-((*S*)-2-(2-(5-chloro-1*H*-indol-2-yl) acetamido)-3-(naphthalen-2-yl) propanamido)-4-ethoxy-4-oxobutanamido) pentanedioate (2.78). ^{1}H NMR (500 MHz, *Chloroform-d*) δ ppm 9.41 (s, 1H), 7.66 (m, 3H), 7.02-7.47 (m, 12H),

6.50 (m, 1H), 6.13 (s, 1H), 4.80 (ddd, 1H, *J* = 6.2 Hz, *J* = 10.5 Hz, *J* = 13.5 Hz), 4.50 (m, 2H), 4.12 (m, 4H), 3.98 (m, 1H), 3.88 (m, 1H), 3.66 (s, 3H), 3.15 (m, 2H), 2.94 (dd, 1H, *J* = 4.6 Hz, *J* = 17.1 Hz), 2.49 (dd, 1H, *J* = 6.0 Hz, *J* = 17.2 Hz), 2.34 (m, 2H), 2.16 (m, 1H), 1.95 (m, 1H), 1.20 (m, 6H). ^{13}C NMR (125 MHz, $CDCl_3$) 172.4, 170.6, 170.6, 135.3, 133.4, 133.1, 132.8, 132.5, 129.5, 128.7, 128.1, 128.0, 127.8, 127.7, 126.9, 126.6, 126.1, 125.6, 122.4, 119.8, 112.3, 102.1, 61.9, 61.4, 61.0, 55.2, 52.3, 49.2, 37.4, 36.3, 35.7, 30.7, 27.3, 14.3, 14.2. HRMS (ESI): *m/e* calcd. for $C_{37}H_{42}ClN_4O_9$ $[M + H]^+$ 722.2038, found 722.2053.



2.49a

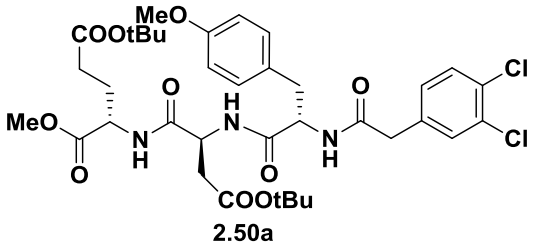
(*S*)-Methyl 2-(2-(3,4-dichlorophenyl) acetamido)-3-(4-methoxyphenyl) propanoate (2.49a). Colorless oil; yield, 90%. ^{1}H NMR (500 MHz, *Chloroform-d*) δ ppm 7.36 (d, 1H, *J* = 8.2 Hz), 7.31 (d, 1H, *J* = 2.0 Hz), 7.03 (dd, 1H, *J* = 2.0 Hz, *J* = 8.2 Hz), 6.82 (d, 2H, *J* = 8.6 Hz), 6.74 (d, 2H, *J* = 8.6 Hz), 6.03 (d, 1H, *J* = 7.8 Hz), 4.80 (td, 1H, *J* = 5.8 Hz, *J* = 7.8 Hz), 3.76 (s, 3H), 3.72 (s, 3H), 3.45 (d, 2H, *J* = 3.1 Hz), 3.02 (dd, 1H, *J* = 5.6 Hz, *J* = 14.1 Hz), 2.95 (dd, 1H, *J* = 6.0 Hz, *J* = 14.1 Hz). ^{13}C NMR (125 MHz, $CDCl_3$) δ ppm 172.1, 169.5, 158.9, 134.9, 132.9, 131.6, 131.4, 130.9, 130.3, 128.9, 127.4, 114.2, 55.4, 53.4, 52.6, 42.6, 36.8.



2.49b

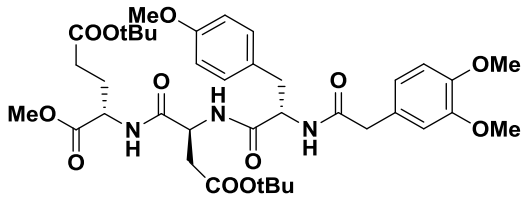
(*S*)-Methyl 2-(2-(3,4-dimethoxyphenyl) acetamido)-3-(4-methoxyphenyl)

propanoate (2.49b). Colorless oil; yield, 93%. ^{1}H NMR (500 MHz, *Chloroform-d*) δ ppm 6.80 (d, 1H, *J* = 8.1 Hz), 6.77 (d, 2H, *J* = 8.5 Hz), 6.71 (m, 4H), 5.87 (d, 1H, *J* = 7.6 Hz), 4.77 (td, 1H, *J* = 5.7 Hz, *J* = 7.6 Hz), 3.87 (s, 3H), 3.80 (s, 3H), 3.76 (s, 3H), 3.69 (s, 3H), 3.48 (s, 3H), 2.98 (dd, 1H, *J* = 5.5 Hz, *J* = 14.0 Hz), 2.93 (dd, 1H, *J* = 5.8 Hz, *J* = 14.0 Hz). ^{13}C NMR (125 MHz, $CDCl_3$) δ ppm 172.1, 171.1, 158.8, 149.5, 148.5, 130.3, 127.5, 127.0, 121.8, 114.1, 112.5, 111.7, 56.1, 56.0, 55.3, 53.2, 52.5, 43.4, 36.9.



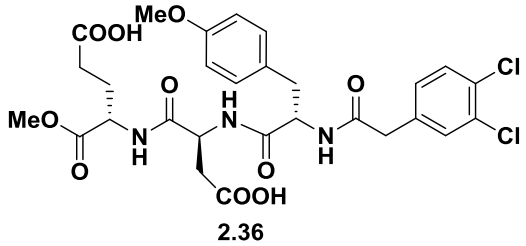
2.50a

(*S*)-5-*tert*-Butyl 1-methyl 2-((*S*)-4-(*tert*-butoxy)-2-((*S*)-2-(2-(3,4-dichlorophenyl) acetamido)-3-(4-methoxyphenyl) propanamido)-4-oxobutanamido) pentanedioate (2.50a). White syrup; yield, 85%. ^{1}H NMR (500 MHz, *Chloroform-d*) δ ppm 7.34 (d, 1H, *J* = 8.2 Hz), 7.30–7.15 (m, 3H), 7.00 (m, 2H), 6.88 (d, 1H, *J* = 8.0 Hz), 6.75 (m, 2H), 6.36 (d, 0.4H, *J* = 6.3 Hz), 6.00 (d, 0.6H, *J* = 6.4 Hz), 4.71 (m, 1H), 4.54–4.40 (m, 2H), 3.77 (s, 3H), 3.69 (d, 2H, *J* = 14.9 Hz), 3.49 (m, 2H), 2.94 (m, 3H), 2.46 (dd, 0.5H, *J* = 6.2 Hz, *J* = 17.3 Hz), 2.28 (td, 2.5H, *J* = 6.3 Hz, *J* = 14.1 Hz), 2.13 (m, 1H), 1.92 (m, 1H), 1.41 (m, 18H). ^{13}C NMR (125 MHz, $CDCl_3$) δ ppm 172.3, 172.2, 172.1, 171.9, 171.6, 171.0, 170.7, 170.5, 170.3, 170.2, 159.0, 158.9, 134.8, 134.6, 133.1, 132.9, 131.9, 131.6, 131.5, 131.1, 130.9, 130.3, 130.2, 129.1, 129.0, 127.9, 127.5, 114.4, 114.3, 82.3, 82.2, 80.9, 77.5, 77.3, 77.0, 55.9, 55.4, 55.0, 52.7, 52.6, 52.1, 52.0, 49.3, 49.2, 42.6, 42.3, 36.8, 36.7, 36.5, 31.6, 28.3, 28.2, 27.4. MS (ESI) *m*/*e* = 752.2 $[M + H]^{+}$.



2.50b

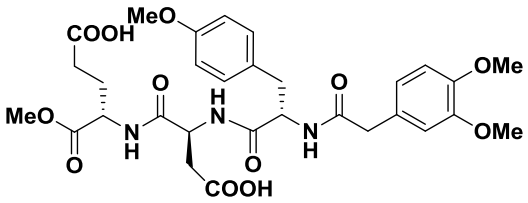
(*S*)-5-*tert*-butyl 1-methyl 2-((*S*)-4-(*tert*-butoxy)-2-((*S*)-2-(2-(3,4-dimethoxyphenyl) acetamido)-3-(4-methoxyphenyl) propanamido)-4-oxobutanamido) pentanedioate (2.50b). White syrup; yield, 87%. ^{1}H NMR (500 MHz, *Chloroform-d*) δ ppm 7.33–7.15 (m, 2H), 6.90 (d, 1H, *J* = 7.0 Hz), 6.80 (m, 2H), 6.70 (m, 3H), 6.60 (s, 1H), 6.00 (d, 0.5H, *J* = 5.9 Hz), 5.85 (d, 0.5H, *J* = 5.7 Hz), 4.80–4.67 (m, 1H), 4.55–4.31 (m, 2H), 3.89 (d, 3H, *J* = 6.8 Hz), 3.78 (m, 6H), 3.70 (m, 3H), 3.48 (m, 2H), 2.92 (m, 3H), 2.48 (m, 1H), 2.26 (m, 3H), 2.12 (m, 1H), 1.95 (m, 1H), 1.43 (m, 18H). ^{13}C NMR (125 MHz, $CDCl_3$) δ ppm 172.5, 172.3, 172.0, 171.7, 171.0, 170.6, 170.4, 170.3, 158.9, 149.6, 149.5, 148.6, 130.2, 130.1, 127.9, 127.6, 126.8, 126.6, 122.0, 121.9, 114.3, 112.6, 111.8, 111.7, 82.2, 82.1, 80.8, 80.7, 56.1, 55.9, 55.3, 52.6, 52.5, 52.1, 49.4, 49.2, 43.3, 37.3, 36.7, 36.4, 31.6, 31.5, 28.3, 28.2, 27.4. MS (ESI) *m*/*e* = 744.4 $[M + H]^+$.



2.36

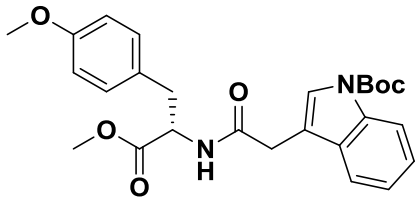
(*S*)-4-((*S*)-3-carboxy-2-((*S*)-2-(2-(3,4-dichlorophenyl) acetamido)-3-(4-methoxyphenyl) propanamido) propanamido)-5-methoxy-5-oxopentanoic acid (2.36). White solid; yield, 90%; m.p. 182.0–184.0 ˚C. ^{1}H NMR (500 MHz, DMSO) δ ppm 12.21 (br, 1H), 8.45 (m, 1H), 8.35 (m, 1H), 8.13 (t, 1H, *J* = 6.7 Hz), 7.45 (m, 1H), 7.36 (d, 1H, *J* = 11.0 Hz), 7.06 (m, 1H), 6.73 (dd, 1H, *J* = 5.7 Hz, *J* = 11.8 Hz), 4.58 (m, 1H), 4.43 (m,

1H), 4.24 (m, 1H), 3.68 (d, 1H, *J* = 1.2 Hz), 3.59 (d, 1H, *J* = 6.9 Hz), 3.45 (dd, 1H, *J* = 8.7 Hz, *J* = 14.0 Hz), 3.33 (m, 1H), 2.89 (m, 1H), 2.63 (m, 1H), 2.41 (m, 1H), 2.23 (m, 1H), 1.92 (dq, 1H, *J* = 7.5 Hz, *J* = 12.9 Hz), 1.76 (ddd, 1H, *J* = 9.1 Hz, *J* = 16.2 Hz, *J* = 24.1 Hz). ^{13}C NMR (125 MHz, DMSO) δ ppm 174.4, 172.6, 172.2, 172.0, 171.8, 171.5, 170.0, 169.7, 168.5, 158.4, 138.0, 131.7, 131.2, 130.9, 130.8, 130.2, 130.1, 130.0, 129.6, 128.9, 114.0, 95.0, 55.5, 55.2, 54.6, 52.6, 52.1, 52.0, 50.1, 49.7, 41.6, 41.4, 37.4, 36.8, 30.4, 26.7, 26.5. HRMS (ESI) *m/e* calcd. for $C_{28}H_{30}Cl_2N_3O_{10}$ $[M - H]^-$ 638.1314, found 638.1335.



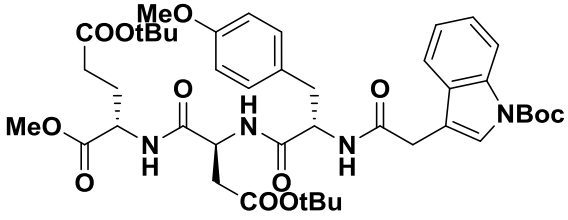
2.37

(*S*)-4-((*S*)-3-carboxy-2-((*S*)-2-(2-(3,4-dimethoxyphenyl) acetamido)-3-(4-methoxyphenyl) propanamido) propanamido)-5-methoxy-5-oxopentanoic acid (2.37). Off-white solid; yield, 96%; m.p. 77.6–79.2 ˚C. ^{1}H NMR (500 MHz, DMSO) δ ppm 12.24 (br, 1H), 8.40 (t, 0.5H, *J* = 8.3 Hz), 8.23–8.06 (m, 1H), 7.23 (t, 0.5H, *J* = 7.5 Hz), 7.16–7.06 (m, 2.5H), 6.74 (m, 3.5H), 6.58 (ddd, 1H, *J* = 1.5 Hz, *J* = 8.1 Hz, *J* = 15.5 Hz), 4.64–4.51 (m, 1H), 4.42–4.18 (m, 2H), 3.68 (m, 6H), 3.65 (m, 3H), 3.59 (d, 3H, *J* = 9.5 Hz), 3.33 (dd, 1H, *J* = 10.8 Hz, *J* = 13.7 Hz), 3.25 (dd, 1H, *J* = 7.7 Hz, *J* = 13.9 Hz), 2.88 (m, 1H), 2.63 (m, 2H), 2.41 (m, 1H), 2.25 (m, 2H), 1.91 (m, 1H), 1.78 (m, 1H). ^{13}C NMR (125 MHz, DMSO) δ ppm 174.4, 172.6, 172.3, 171.9, 171.5, 171.0, 158.4, 149.0, 148.0, 138.0, 130.9, 130.3, 130.0, 129.6, 129.2, 128.9, 126.0, 121.7, 121.6, 114.0, 113.5, 112.3, 56.1, 55.5, 52.6, 52.5, 52.1, 50.1, 42.4, 42.3, 37.3, 36.8, 36.7, 30.4, 26.7. HRMS (ESI) *m/e* calcd. for $C_{30}H_{36}N_3O_{12}$ $[M - H]^-$ 630.2304, found 630.2309.



2.60b

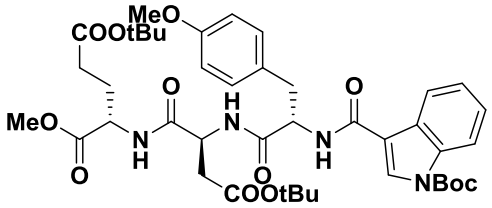
(*S*)-*tert*-Butyl 3-(2-((1-methoxy-3-(4-methoxyphenyl)-1-oxopropan-2-yl) amino)-2-oxoethyl)-1*H*-indole-1-carboxylate (2.60b). Colorless oil; yield, 95%. ^{1}H NMR (300 MHz, *Chloroform-d*) δ ppm 8.21 (d, 1H, *J* = 7.8 Hz), 7.55 (s, 1H), 7.49 (d, 1H, *J* = 7.7 Hz), 7.39 (dt, 1H, *J* = 1.1 Hz, *J* = 7.8 Hz), 7.27 (t, 1H, *J* = 7.5 Hz), 6.66 (d, 2H, *J* = 8.7 Hz), 6.59 (d, 2H, *J* = 8.8 Hz), 6.12 (d, 1H, *J* = 8.0 Hz), 4.85 (td, 1H, *J* = 5.6 Hz, *J* = 8.0 Hz), 3.74 (s, 3H), 3.70 (s, 3H), 3.67 (s, 2H), 2.95 (d, 2H, *J* = 5.7 Hz), 1.71 (s, 9H). ^{13}C NMR (75 MHz, $CDCl_3$) δ ppm 171.9, 169.e8, 158.6, 149.5, 135.7, 130.0, 129.8, 127.2, 124.9, 122.9, 119.2, 115.4, 113.8, 113.5, 83.9, 55.1, 53.0, 52.3, 36.7, 33.2, 28.2.



2.61b

(*S*)-5-*tert*-Butyl 1-methyl 2-((*S*)-4-(*tert*-butoxy)-2-((*S*)-2-(2-(1-(*tert*-butoxycarbonyl)-1*H*-indol-3-yl) acetamido)-3-(4-methoxyphenyl) propanamido)-4-oxobutanamido) pentanedioate (2.61b). Light orange oil; yield, 87%. ^{1}H NMR (500 MHz) 8.97 (s, 0.5H), 8.69 (s, 0.5H), 7.49 (d, 0.5H, *J* = 8.0 Hz), 7.46 (d, 0.5H, *J* = 7.9 Hz), 7.40 (m, 1H), 7.31 (t, 1H, *J* = 8.8 Hz), 7.23 (m, 2H), 7.11 (m 1.5H), 7.01 (m, 0.5H), 6.90 (d, 0.5H, *J* = 8.1 Hz), 6.68 (d, 1H, *J* = 8.6 Hz), 6.52 (m, 3H), 6.28 (m, 0.5H), 6.17 (d, 0.5H, *J* = 6.0 Hz), 4.80 (dt, 0.5H, *J* = 4.4 Hz, *J* = 7.3 Hz), 4.69 (m, 1H), 4.52 (m, 1.5H), 4.45 (m,

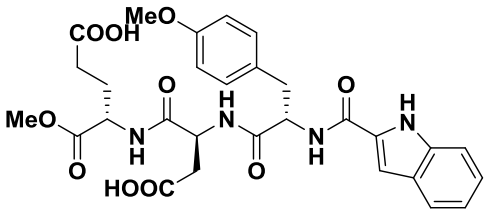
1H), 3.70 (m, 8H), 3.06–2.75 (m, 3H), 2.77 (dd, 0.5H, *J* = 6.8 Hz, *J* = 13.8 Hz), 2.57 (m, 0.5H), 2.30 (m, 2H), 2.14 (m, 1H), 1.91 (m, 1H), 1.42 (m, 27H). ^{13}C NMR (75 MHz, $CDCl_3$) δ ppm 173.0, 172.6, 172.6, 172.4, 172.3, 172.2, 172.0, 171.9, 171.8, 171.7, 171.2, 170.9, 170.7, 170.5, 170.4, 170.2, 167.5, 167.3, 158.6, 158.6, 136.7, 136.6, 130.2, 130.0, 127.9, 127.6, 127.2, 127.1, 124.4, 124.3, 122.8, 122.6, 120.4, 120.2, 118.7, 114.2, 114.1, 111.7, 108.2, 108.1, 99.2, 82.5, 82.1, 81.4, 81.1, 81.1, 80.8, 55.4, 55.3, 55.3, 54.6, 52.6, 52.2, 52.1, 51.9, 49.4, 49.2, 39.4, 37.3, 36.7, 36.5, 36.2, 33.2, 33.1, 32.1, 31.8, 31.6, 31.5, 31.3, 29.8, 29.6, 28.3, 28.2, 27.5, 27.4, 27.3, 23.4, 22.9. MS (ESI) *m*/*e* = 822.5 $[M + H]^+$.



2.61c

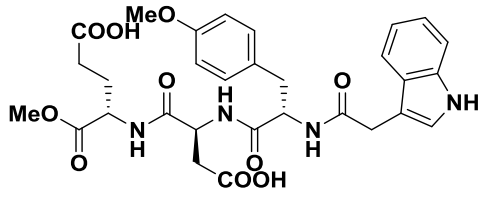
(*S*)-5-*tert*-Butyl 1-methyl 2-((*S*)-2-((*S*)-2-(1*H*-indole-3-carboxamido)-3-(4-methoxyphenyl) propanamido)-4-(*tert*-butoxy)-4-oxobutanamido) pentanedioate (2.61c). Light orange syrup; yield, 85%. ^{1}H NMR (500 MHz, *Chloroform-d*) δ ppm 9.83 (s, 1H), 7.75 (d, 1H, *J* = 2.9 Hz), 7.66 (d, 1H, *J* = 7.8 Hz), 7.59 (d, 1H, *J* = 8.4 Hz), 7.43 (d, 1H, *J* = 8.0 Hz), 7.32 (d, 1H, *J* = 8.1 Hz), 7.12 (m, 5H), 6.86 (d, 1H, *J* = 6.0 Hz), 6.79 (d, 2H, *J* = 8.6 Hz), 4.80 (m, 2H), 4.50 (dt, 1H, *J* = 5.3 Hz, *J* = 8.2 Hz), 3.70 (s, 3H), 3.65 (s, 3H), 3.21 (dd, 1H, *J* = 6.0 Hz, *J* = 14.0 Hz), 3.14 (dd, 1H, *J* = 7.6 Hz, *J* = 14.0 Hz), 3.00 (dd, 1H, *J* = 3.2 Hz, *J* = 17.3 Hz), 2.90 (dd, 1H, *J* = 4.7 Hz, *J* = 17.2 Hz), 2.60 (ddd, 1H, *J* = 7.9 Hz, *J* = 13.1 Hz, *J* = 17.2 Hz), 2.38 (t, 1H, *J* = 7.3 Hz), 2.24 (m, 2H), 2.14 (m, 1H), 1.92 (m, 1H), 1.36 (s, 9H), 1.31 (s, 9H). ^{13}C NMR (125 MHz, $CDCl_3$) δ ppm 172.4, 172.1, 171.8, 171.4, 170.6, 166.3, 159.0, 136.5, 130.5, 129.6, 128.5, 124.8, 123.0, 121.8, 120.0, 114.6, 112.3, 111.2, 82.0, 80.8, 55.6, 55.4, 52.6, 52.2, 49.5, 39.5, 37.0, 31.6, 28.3,

28.2, 27.3. MS (ESI) *m/e* = 709.3 $[M + H]^+$.



2.38

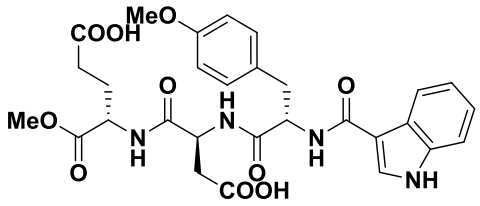
(*S*)-4-((*S*)-2-((*S*)-2-(1*H*-Indole-2-carboxamido)-3-(4-methoxyphenyl) propanamido)-3-carboxypropanamido)-5-methoxy-5-oxopentanoic acid (2.38). Light yellow solid; yield, quantitative; m.p. decompose at 107 ˚C. ^{1}H NMR (500 MHz, DMSO) δ ppm 12.24 (br, 2H), 11.39 (s, 1H), 8.52 (ddd, 2H, *J* = 8.6 Hz, *J* = 15.5 Hz, *J* = 26.1 Hz), 8.19 (dd, 1H, *J* = 7.3 Hz, *J* = 18.5 Hz), 7.60 (d, 3H, *J* = 7.4 Hz), 7.38 (t, 1H, *J* = 7.7 Hz), 7.24 (t, 2H, *J* = 8.1 Hz), 7.19 (s, 1H), 7.15 (m, 1H), 7.01 (m, 1H), 6.78 (dd, 2H, *J* = 2.3 Hz, *J* = 8.7 Hz), 4.64 (m, 2H), 4.26 (m, 1H), 3.64 (s, 3H), 3.60 (s, 3H), 3.57 (m, 2H), 3.03 (m, 1H), 2.88 (m, 1H), 2.66 (dt, 1H, *J* = 5.1 Hz, *J* = 16.7 Hz), 2.27 (m, 2H), 1.96 (m, 1H), 1.83 (m, 1H). HRMS (ESI) *m/e* calcd. for $C_{29}H_{31}N_4O_{10}$ $[M - H^+]^-$ 595.2046, found 595.2056.



2.39

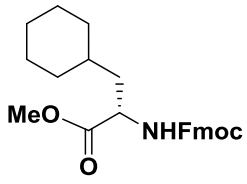
(*S*)-4-((*S*)-2-((*S*)-2-(2-(1*H*-indol-3-yl) acetamido)-3-(4-methoxyphenyl) propanamido)-3-carboxypropanamido)-5-methoxy-5-oxopentanoic acid (2.39). Light brown solid; yield, 96%; m.p. decompose at 92 ˚C. ^{1}H NMR (500 MHz, *Chloroform-d*) δ ppm 12.22 (br, 2H), 10.80 (s, 1H), 8.42 (m, 1H), 8.14 (m, 2H), 6.62 – 7.38 (m, 7H), 4.57 (m, 1H), 4.22 (m, 2H), 3.44 – 3.70 (m, 8H), 2.87 (dt, 0.5H, *J* = 4.6Hz, *J* = 13.7Hz), 2.62

(m, 1.5H) 2.38 (m, 1H) 2.25 (m, 2H) 1.93 (m, 1H) 1.75 (m, 1H) ^{13}C NMR (125 MHz, DMSO) δ ppm 174.9, 174.4, 174.4, 172.7, 172.6, 172.6, 172.3, 172.3, 172.1, 171.8, 171.4, 170.0, 168.4, 168.3, 158.4, 136.7, 130.9, 130.2, 129.9, 127.8, 124.5, 119.4, 118.9, 114.0, 111.8, 109.2, 55.5, 52.6, 52.0, 51.5, 49.9, 36.9, 36.7, 30.4, 26.5. HRMS (ESI) *m/e* calcd. for $C_{30}H_{33}N_4O_{10}$ $[M - H]^-$ 609.2197, found 609.2222.



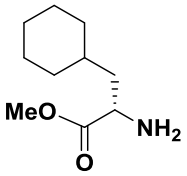
2.40

(*S*)-4-((*S*)-2-((*S*)-2-(1*H*-Indole-3-carboxamido)-3-(4-methoxyphenyl) propanamido)-3-carboxypropanamido)-5-methoxy-5-oxopentanoic acid (2.40). Yellow solid; yield, 95%; m.p. decompose at 104˚C. ^{1}H NMR (500 MHz, DMSO) δ ppm 12.23 (br, 2H), 11.53 (d, 1H, *J* = 2.4 Hz), 8.44 (d, 1H, *J* = 7.8 Hz), 8.11 (m, 3H), 7.97 (d, 1H, *J* = 7.8 Hz), 7.91 (d, 1H, *J* = 7.9 Hz), 7.39 (d, 1H, *J* = 8.1 Hz), 7.23 (t, 3H, *J* = 8.4 Hz), 7.12 (ddd, 3H, *J* = 4.5 Hz, *J* = 13.2 Hz, *J* = 15.1 Hz), 7.04 (m, 1H), 6.78 (d, 3H, *J* = 8.7 Hz), 4.59 (m, 2H), 4.28 (m, 1H), 4.15 (t, 0.5H, *J* = 4.8 Hz), 3.95 (t, 0.5H, *J* = 5.4 Hz), 3.65 (s, 3H), 3.60 (s, 3H), 3.03 (dd, 1H, *J* = 3.6 Hz, *J* = 13.8 Hz), 2.87 (dd, 1H, *J* = 10.4 Hz, *J* = 13.7 Hz), 2.70 (dd, 1H, *J* = 5.3 Hz, *J* = 16.6 Hz), 2.63 (dd, 1H, *J* = 3.0 Hz, *J* = 5.2 Hz), 2.54 (dd, 1H, *J* = 8.0 Hz, *J* = 16.6 Hz), 2.26 (m, 1H), 1.94 (m, 1H), 1.78 (m, 1H). ^{13}C NMR (125 MHz, DMSO) δ ppm 179.8, 179.4, 177.8, 177.6, 177.4, 177.1, 176.5, 173.3, 170.3, 163.4, 141.7, 136.0, 135.9, 134.6, 134.1, 133.9, 131.6, 131.0, 127.5, 126.5, 126.0, 119.1, 117.5, 115.7, 95.0, 60.5, 60.1, 58.8, 57.6, 57.0, 56.5, 55.2, 42.0, 41.8, 41.6, 35.3, 35.0, 32.8, 31.8, 26.7. HRMS (ESI) *m/e* calcd. for $C_{29}H_{31}N_4O_{10}$ $[M - H]^-$ 595.2046, found 595.2057.



2.71c

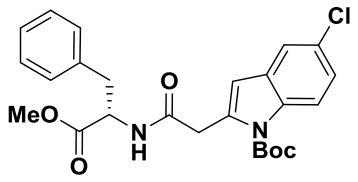
Fmoc-*L*-Hex-OMe (2.71c). Colorless oil; yield, 95%. ^{1}H NMR (300 MHz, *Chloroform-d*) δ ppm 7.76 (d, 2H, *J* = 7.6 Hz), 7.60 (m, 2H), 7.40 (t, 2H, *J* = 7.5 Hz), 7.31 (t, 2H, *J* = 7.2 Hz), 5.18 (d, 1H, *J* = 8.8 Hz), 4.41 (m, 3H), 4.24 (t, 1H, *J* = 7.1 Hz), 3.74 (s, 3H), 1.84–0.83(m, 13H). ^{13}C NMR (75 MHz, $CDCl_3$) δ ppm 174.0, 156.2, 144.0, 141.5, 127.9, 127.3, 125.3, 120.2, 67.2, 52.6, 52.0, 47.4, 40.4, 34.2, 33.8, 32.6, 26.6, 26.4, 26.2.



2.72c

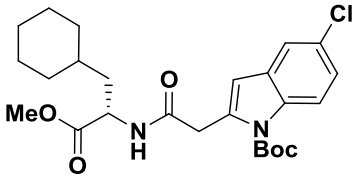
H-L-Hex-OMe (2.72c). Colorless oil; yield, 87%. ^{1}H NMR (300 MHz, *Chloroform-d*) δ ppm 3.66 (s, 3H), 3.46 (dd, 1H, *J* = 5.3 Hz, *J* = 8.5 Hz), 1.73–0.78 (m, 13H). ^{13}C NMR (75 MHz, $CDCl_3$) δ ppm 177.4, 52.3, 52.1, 42.8, 34.2, 33.9, 32.6, 26.7, 26.4, 26.2.

H-*L*-Hex-OMe hydrochloride salt (2.69c). At 0 ˚C, to a stirred solution of H-L-Hex-OMe (**2.72c**) in anhydrous dichloromethane was added 1.2 equiv of 1 M HCl in MeOH. After 10 min, the solvent was evaporated under reduced pressure to yield the desired H-*L*-Hex-OMe hydrochloride salt as a white to off-white solid with a quantitative yield.



2.76b

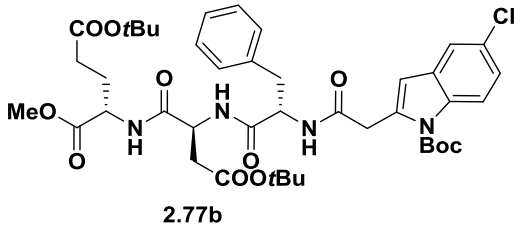
(*S*)-*tert*-Butyl 5-chloro-2-(2-((1-methoxy-1-oxo-3-phenylpropan-2-yl) amino)-2-oxoethyl)-1*H*-indole-1-carboxylate (2.76b). Colorless oil; yield, 96%. ^{1}H NMR (500 MHz, *Chloroform-d*) δ ppm 8.00 (d, 1H, *J* = 8.9 Hz), 7.45 (d, 1H, *J* = 2.1 Hz), 7.25 (dd, 1H, *J* = 2.1 Hz, *J* = 8.9 Hz), 7.11 (t, 1H, *J* = 6.8 Hz), 7.02 (t, 2H, *J* = 7.6 Hz), 6.88 (d, 2H, *J* = 7.7 Hz), 6.44 (s, 1H), 6.32 (d, 1H, *J* = 7.7 Hz), 4.84 (td, 1H, *J* = 5.8 Hz, *J* = 7.7 Hz), 3.94 (dd, 2H, *J* = 15.8 Hz, *J* = 42.2 Hz), 3.68 (s, 3H), 3.05 (dd, 1H, *J* = 5.6 Hz, *J* = 13.8 Hz), 3.01 (dd, 1H, *J* = 5.8 Hz, *J* = 13.8 Hz), 1.61 (s, 9H). ^{13}C NMR (125 MHz, $CDCl_3$) δ ppm 171.8, 168.8, 150.4, 135.8, 135.5, 135.1, 130.2, 129.2, 128.8, 128.7, 127.2, 124.5, 120.1, 117.1, 110.4, 85.6, 53.3, 52.5, 38.2, 37.9, 28.3.



2.76c

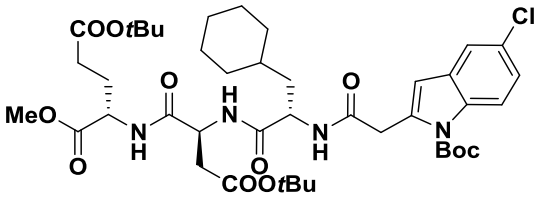
(*S*)-*tert*-Butyl 5-chloro-2-(2-((3-cyclohexyl-1-methoxy-1-oxopropan-2-yl)amino)-2-oxoethyl)-1*H*-indole-1-carboxylate (2.76c). Colorless oil; yield, 80%. ^{1}H NMR (300 MHz, *Chloroform-d*) δ ppm 7.99 (d, 1H, *J* = 8.9 Hz), 7.46 (dd, 1H, *J* = 0.4 Hz, *J* = 2.1 Hz), 7.22 (dd, 1H, *J* = 2.2 Hz, *J* = 8.9 Hz), 6.52 (d, 1H, *J* = 0.6 Hz), 6.18 (d, 1H, *J* = 8.2 Hz), 4.62 (dt, 1H, *J* = 5.4 Hz, *J* = 8.8 Hz), 3.97 (s, 2H), 3.67 (s, 3H), 1.67 (s, 9H), 1.63–0.75 (m, 13H). ^{13}C NMR (75 MHz, $CDCl_3$) δ ppm 173.6, 169.1, 150.4, 135.6, 135.2, 130.2, 128.8, 124.5, 120.1, 117.1, 110.5, 85.5, 52.5, 50.5, 40.2, 38.3, 34.3, 33.6, 32.7, 28.4,

26.5, 26.3, 26.2.



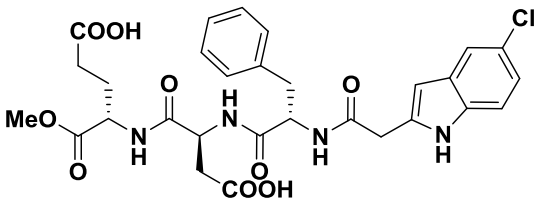
2.77b

(*S*)-5-*tert*-butyl 1-methyl 2-((*S*)-4-(*tert*-butoxy)-2-((*S*)-2-(2-(1-(*tert*-butoxycarbonyl)-5-chloro-1*H*-indol-2-yl) acetamido)-3-phenylpropanamido)-4-oxobutanamido) pentanedioate (2.77b). Light orange oil; yield, 86%. Two rotamers. ^{1}H NMR (500 MHz) ppm 7.95 (t, 1H, *J* = 8.4 Hz, major + minor), 7.45 (d, 1H, *J* = 9.5 Hz, major + minor), 7.26 (m, 2H, major + minor), 7.10 (m, 4H, major + minor), 6.98 (m, 2H, major + minor), 6.92 (s, 0.6H, major), 6.91 (s, 0.4H, minor), 6.46 (s, 0.4H, minor), 6.42 (s, 0.6H, major) 6.38 (d, 0.4H, *J* = 6.2 Hz, minor), 6.23 (d, 0.6H, *J* = 6.0 Hz, major), 4.70 (m, 1H, major + minor), 4.49 (m, 1.6H, major + minor), 4.41 (q, 0.4H, *J* = 6.7 Hz, minor), 3.91 (m, 2H, major + minor), 3.70 (s, 1.6H, major), 3.67 (s, 1.4H, minor), 3.00 (m, 2H, major + minor), 2.88 (m, 1H, major + minor), 2.48 (dd, 0.4H, *J* = 6.4 Hz, *J* = 17.3 Hz, major + minor), 2.27 (m, 2.6H, major + minor), 2.11 (m, 1H, major + minor), 1.92 (m, 1H, major + minor), 1.60 (m, 9H, major + minor), 1.41 (s, 18H, major + minor). ^{13}C NMR (125 MHz, $CDCl_3$) δ ppm 172.2, 172.1, 172.0, 171.9, 171.7, 171.6, 170.7, 170.5, 170.4, 170.3, 170.0, 150.3, 136.1, 135.9, 135.1, 135.1, 135.0, 134.8, 130.1, 130.1, 129.1, 129.0, 128.9, 128.8, 127.3, 124.7, 124.5, 120.1, 117.1, 110.8, 110.6, 95.0, 85.7, 85.5, 82.1, 82.0, 80.8, 55.9, 55.1, 52.6, 52.5, 52.2, 52.1, 49.3, 49.2, 38.1, 38.0, 37.5, 37.0, 36.6, 31.6, 28.3, 28.2, 27.3.



2.77c

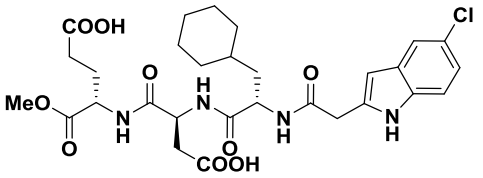
(*S*)-5-*tert*-Butyl 1-methyl 2-((*S*)-4-(*tert*-butoxy)-2-((*S*)-2-(2-(1-(*tert*-butoxycarbonyl)-5-chloro-1*H*-indol-2-yl) acetamido)-3-cyclohexylpropanamido)-4-oxobutanamido) pentanedioate (2.77c). Orange syrup; yield, 72%. ^{1}H NMR (500 MHz, *Chloroform-d*) δ ppm 7.94 (d, 0.5H, *J* = 3.5 Hz), 7.93 (d, 0.5H, *J* = 3.5 Hz), 7.46 (s, 1H), 7.38 (d, 0.5H, *J* = 8.0 Hz), 7.27 (d, 0.5H, *J* = 8.9 Hz), 7.25 (d, 0.5H, *J* = 3.5 Hz), 7.22 (m, 1.5H), 6.54 (d, 1H, *J* = 6.5 Hz), 6.29 (dd, 1H, *J* = 2.6 Hz, *J* = 6.5 Hz), 4.76 (m, 0.5H), 4.72 (m, 0.5H), 4.50 (dq, 1H, *J* = 5.1 Hz, *J* = 8.0 Hz), 4.37 (td, 0.5H, *J* = 5.9 Hz, *J* = 9.7 Hz), 4.26 (dd, 0.5H, *J* = 6.6 Hz, *J* = 14.9 Hz), 3.97 (m, 2H), 3.70 (s, 1.7H), 3.68 (s, 1.3H), 3.00 (dd, 0.4H, *J* = 4.2 Hz, *J* = 17.3 Hz), 2.90 (dd, 0.6H, *J* = 4.3 Hz, *J* = 17.2 Hz), 2.53 (m, 1H), 2.26 (m, 1H), 2.11 (m, 2H), 1.93 (m, 1H), 1.67 (s, 9H), 1.54–1.63 (m, 5H), 1.42 (m, 18H), 0.79–1.31 (m, 8H). ^{13}C NMR (125 MHz, $CDCl_3$) δ ppm 172.2, 172.0, 171.9, 171.7, 170.5, 170.1, 150.6, 150.5, 135.4, 135.3, 135.0, 134.9, 130.2, 128.9, 128.8, 124.6, 124.5, 120.2, 120.1, 117.0, 110.8, 110.7, 85.7, 85.5, 82.1, 82.0, 80.8, 52.6, 52.4, 52.2, 52.1, 49.3, 39.5, 38.8, 38.1, 37.9, 37.0, 36.8, 34.4, 33.8, 33.5, 32.9, 32.6, 31.6, 31.5, 29.9, 28.4, 28.2, 27.4, 27.3, 26.2. MS (ESI) *m*/*e* = 834.4 $[M + H]^+$.



2.63

(*S*)-4-((*S*)-3-carboxy-2-((*S*)-2-(2-(5-chloro-1*H*-indol-2-yl) acetamido)-3-

phenylpropanamido) propanamido)-5-methoxy-5-oxopentanoic acid (2.63). Pink solid; yield, quantitative; m.p. decompose at 103˚C. ^{1}H NMR (500 MHz, DMSO) δ ppm 12.24 (br, 2H), 11.03 (s, 1H), 8.47 (dd, 1H, *J* = 4.3 Hz, *J* = 7.9 Hz), 8.31 (dd, 1H, *J* = 7.6 Hz, *J* = 21.2 Hz), 8.18 (dd, 1H, *J* = 7.6 Hz, *J* = 12.6 Hz), 7.41 (s, 1H), 7.28 (d, 1H, *J* = 8.6 Hz), 7.23 (t, 1H, *J* = 7.5 Hz), 7.15 (m, 8H), 6.97 (dd, 1H, *J* = 1.9 Hz, *J* = 8.6 Hz), 6.02 (dd, 1H, *J* = 0.9 Hz, *J* = 18.1 Hz), 4.62 (m, 0.5H), 4.53 (m, 2H), 4.23 (m, 1.5H), 3.60 (s, 3H), 3.58 (s, 2H), 3.53 (m, 2H), 2.98 (ddd, 1H, *J* = 4.4 Hz, *J* = 13.8 Hz, *J* = 18.8 Hz), 2.74 (m, 1H), 2.63 (ddd, 1H, *J* = 5.0 Hz, *J* = 16.6 Hz, *J* = 31.2 Hz), 2.39 (dd, 1H, *J* = 8.3 Hz, *J* = 16.5 Hz), 2.25 (m, 4H), 1.91 (m, 1H), 1.77 (m, 1H). ^{13}C NMR (125 MHz, DMSO) δ ppm 174.4, 172.7, 172.2, 171.9, 171.6, 171.5, 169.2, 169.1, 138.4, 138.1, 136.3, 136.2, 135.1, 129.9, 129.9, 129.6, 128.9, 128.6, 128.6, 126.8, 126.0, 123.8, 120.8, 119.0, 112.9, 100.4, 100.4, 54.9, 54.6, 52.6, 52.5, 52.1, 52.0, 50.2, 49.8, 35.8, 30.5, 30.4, 26.7. HRMS (ESI) *m/e* calcd. for $C_{29}H_{30}ClN_4O_9$ $[M - H]^-$ 613.1707, found 613.1712.



2.64

(*S*)-4-((*S*)-3-Carboxy-2-((*S*)-2-(2-(5-chloro-1*H*-indol-2-yl) acetamido)-3-cyclohexyl propanamido) propanamido)-5-methoxy-5-oxopentanoic acid (2.64). Pink solid; yield, 98%; m.p. decompose at 82 ˚C. ^{1}H NMR (500 MHz, DMSO) δ ppm 11.07 (s, 0.6H), 11.01 (s, 0.4H), 9.52 (s, 1H), 8.32 (m, 1H), 8.05 (dd, 1H, *J* = 7.6 Hz, *J* = 27.0 Hz), 7.42 (t, 1H, *J* = 2.6 Hz), 7.29 (dd, 1H, *J* = 3.8 Hz, *J* = 8.6 Hz), 7.22 (t, 1H, *J* = 7.3 Hz), 7.13 (dd, 1H, *J* = 7.4 Hz, *J* = 18.3 Hz), 6.97 (td, 1H, *J* = 2.1 Hz, *J* = 8.5 Hz), 6.19 (s, 1H), 4.58 (m, 0.5H), 4.49 (m, 1H), 4.23 (m, 1.5H), 3.58 (m, 2H), 3.55 (s, 3H), 2.67 (ddd, H, *J*

= 4.8 Hz, *J* = 9.4 Hz, *J* = 16.6 Hz), 2.44 (m, 1H), 2.21 (m, 2H), 1.88 (m, 1H), 1.74 (m, 1H), 0.71–1.63 (m, 13H). ^{13}C NMR (125 MHz, DMSO) δ ppm 174.5, 174.4, 173.0, 172.6, 172.3, 172.2, 171.6, 171.4, 169.6, 169.4, 136.5, 136.4, 135.2, 135.1, 129.9, 129.6, 128.9, 126.0, 123.9, 120.9, 119.0, 113.0, 112.9, 100.4, 52.6, 52.5, 52.1, 52.0, 51.8, 51.2, 50.1, 49.9, 36.4, 35.9, 35.7, 34.0, 33.9, 33.8, 32.4, 30.4, 30.3, 30.2, 26.7, 26.4, 26.2, 26.1. HRMS (ESI) *m/e* calcd. for $C_{29}H_{36}ClN_4O_9$ $[M - H]^-$ 619.2176, found 619.2181.

2.5 References

1. Clevers, H.; Nusse, R., Wnt/b-catenin signaling and disease. *Cell* **2012,** *149* (6), 1192-1205.

2. Malanchi, I.; Peinado, H.; Kassen, D.; Hussenet, T.; Metzger, D.; Chambon, P.; Huber, M.; Hohl, D.; Cano, A.; Birchmeier, W.; Huelsken, J., Cutaneous cancer stem cell maintenance is dependent on beta-catenin signalling. *Nature* **2008,** *452* (7187), 650-3.

3. Barker, N.; Ridgway, R. A.; van Es, J. H.; van de Wetering, M.; Begthel, H.; van den Born, M.; Danenberg, E.; Clarke, A. R.; Sansom, O. J.; Clevers, H., Crypt stem cells as the cells-of-origin of intestinal cancer. *Nature* **2009,** *457* (7229), 608-11.

4. Yeung, J.; Esposito, M. T.; Gandillet, A.; Zeisig, B. B.; Griessinger, E.; Bonnet, D.; So, C. W., beta-Catenin mediates the establishment and drug resistance of MLL leukemic stem cells. *Cancer Cell* **2010,** *18* (6), 606-18.

5. Barker, N.; Clevers, H., Mining the Wnt pathway for cancer therapeutics. *Nature reviews. Drug Discovery* **2006,** *5* (12), 997-1014.

6. Polakis, P., Drugging Wnt signalling in cancer. *EMBO J* **2012,** *31* (12), 2737-46.

7. Anastas, J. N.; Moon, R. T., WNT signalling pathways as therapeutic targets in cancer. *Nature Reviews. Cancer* **2013,** *13* (1), 11-26.

8. Weber, B. N.; Chi, A. W.; Chavez, A.; Yashiro-Ohtani, Y.; Yang, Q.; Shestova, O.; Bhandoola, A., A critical role for TCF-1 in T-lineage specification and differentiation. *Nature* **2011,** *476* (7358), 63-8.

9. Graham, T. A.; Weaver, C.; Mao, F.; Kimelman, D.; Xu, W., Crystal structure of a beta-catenin/Tcf complex. *Cell* **2000,** *103* (6), 885-96.

10. Poy, F.; Lepourcelet, M.; Shivdasani, R. A.; Eck, M. J., Structure of a human Tcf4-beta-catenin complex. *Nature Structural Biology* **2001,** *8* (12), 1053-7.

11. Graham, T. A.; Ferkey, D. M.; Mao, F.; Kimelman, D.; Xu, W., Tcf4 can specifically recognize beta-catenin using alternative conformations. *Nature Structural Biology* **2001,** *8* (12), 1048-52.

12. Sampietro, J.; Dahlberg, C. L.; Cho, U. S.; Hinds, T. R.; Kimelman, D.; Xu, W., Crystal structure of a beta-catenin/BCL9/Tcf4 complex. *Molecular Cell* **2006,** *24* (2), 293-300.

13. Sun, J.; Weis, W. I., Biochemical and structural characterization of beta-catenin interactions with nonphosphorylated and CK2-phosphorylated Lef-1. *Journal of Molecular Biology* **2011,** *405* (2), 519-30.

14. Huber, A. H.; Weis, W. I., The structure of the beta-catenin/E-cadherin complex and the molecular basis of diverse ligand recognition by beta-catenin. *Cell* **2001,** *105* (3), 391-402.

15. Ha, N. C.; Tonozuka, T.; Stamos, J. L.; Choi, H. J.; Weis, W. I., Mechanism of phosphorylation-dependent binding of APC to beta-catenin and its role in beta-catenin degradation. *Molecular Cell* **2004,** *15* (4), 511-21.

16. Xing, Y.; Clements, W. K.; Le Trong, I.; Hinds, T. R.; Stenkamp, R.; Kimelman, D.; Xu, W., Crystal structure of a beta-catenin/APC complex reveals a critical role for APC phosphorylation in APC function. *Molecular Cell* **2004,** *15* (4), 523-33.

17. Hulsken, J.; Birchmeier, W.; Behrens, J., E-cadherin and APC compete for the interaction with beta-catenin and the cytoskeleton. *The Journal of Cell Biology* **1994,** *127* (6 Pt 2), 2061-9.

18. Omer, C. A.; Miller, P. J.; Diehl, R. E.; Kral, A. M., Identification of Tcf4 residues involved in high-affinity beta-catenin binding. *Biochemical and Biophysical Research Communications* **1999,** *256* (3), 584-90.

19. Orsulic, S.; Huber, O.; Aberle, H.; Arnold, S.; Kemler, R., E-cadherin binding prevents beta-catenin nuclear localization and beta-catenin/LEF-1-mediated transactivation. *Journal of Cell Science* **1999,** *112 (Pt 8)*, 1237-45.

20. Choi, H. J.; Huber, A. H.; Weis, W. I., Thermodynamics of beta-catenin-ligand interactions: the roles of the N- and C-terminal tails in modulating binding affinity. *The Journal of Biological Chemistry* **2006,** *281* (2), 1027-38.

21. Knapp, S.; Zamai, M.; Volpi, D.; Nardese, V.; Avanzi, N.; Breton, J.; Plyte, S.; Flocco, M.; Marconi, M.; Isacchi, A.; Caiolfa, V. R., Thermodynamics of the high-affinity interaction of TCF4 with beta-catenin. *Journal of Molecular Biology* **2001,** *306* (5), 1179-89.

22. Lepourcelet, M.; Chen, Y. N.; France, D. S.; Wang, H.; Crews, P.; Petersen, F.; Bruseo, C.; Wood, A. W.; Shivdasani, R. A., Small-molecule antagonists of the oncogenic Tcf/beta-catenin protein complex. *Cancer Cell* **2004,** *5* (1), 91-102.

23. Trosset, J. Y.; Dalvit, C.; Knapp, S.; Fasolini, M.; Veronesi, M.; Mantegani, S.; Gianellini, L. M.; Catana, C.; Sundstrom, M.; Stouten, P. F.; Moll, J. K., Inhibition of protein-protein interactions: the discovery of druglike beta-catenin inhibitors by combining virtual and biophysical screening. *Proteins* **2006,** *64* (1), 60-7.

24. Gonsalves, F. C.; Klein, K.; Carson, B. B.; Katz, S.; Ekas, L. A.; Evans, S.; Nagourney, R.; Cardozo, T.; Brown, A. M.; DasGupta, R., An RNAi-based chemical genetic screen identifies three small-molecule inhibitors of the Wnt/wingless signaling pathway. *Proceedings of the National Academy of Sciences of the United States of America* **2011,** *108* (15), 5954-63.

25. Tian, W.; Han, X.; Yan, M.; Xu, Y.; Duggineni, S.; Lin, N.; Luo, G.; Li, Y. M.; Han, X.; Huang, Z.; An, J., Structure-based discovery of a novel inhibitor targeting the beta-catenin/Tcf4 interaction. *Biochemistry* **2012,** *51* (2), 724-31.

26. Zhang, M.; Catrow, J. L.; Ji, H., High-throughput selectivity assays for small-molecule inhibitors of beta-catenin/T-cell factor protein-protein interactions. *ACS Medicinal Chemistry Letters* **2013,** *4* (2), 306-11.

27. Grossmann, T. N.; Yeh, J. T.; Bowman, B. R.; Chu, Q.; Moellering, R. E.; Verdine, G. L., Inhibition of oncogenic Wnt signaling through direct targeting of beta-catenin. *Proceedings of the National Academy of Sciences of the United States of America* **2012,** *109* (44), 17942-7.

28. Yu, B.; Huang, Z.; Zhang, M.; Dillard, D. R.; Ji, H., Rational design of small-molecule inhibitors for beta-catenin/T-cell factor protein-protein interactions by bioisostere replacement. *ACS Chemical Biology* **2013,** *8* (3), 524-9.

29. Fasolini, M.; Wu, X.; Flocco, M.; Trosset, J. Y.; Oppermann, U.; Knapp, S., Hot spots in Tcf4 for the interaction with beta-catenin. *The Journal of Biological Chemistry* **2003,** *278* (23), 21092-8.

30. Gail, R.; Frank, R.; Wittinghofer, A., Systematic peptide array-based delineation of the differential beta-catenin interaction with Tcf4, E-cadherin, and adenomatous polyposis coli. *The Journal of Biological Chemistry* **2005,** *280* (8), 7107-17.

31. von Kries, J. P.; Winbeck, G.; Asbrand, C.; Schwarz-Romond, T.; Sochnikova, N.; Dell'Oro, A.; Behrens, J.; Birchmeier, W., Hot spots in beta-catenin for interactions with LEF-1, conductin and APC. *Nature Structural Biology* **2000,** *7* (9), 800-7.

32. Wells, J. A.; McClendon, C. L., Reaching for high-hanging fruit in drug discovery at protein-protein interfaces. *Nature* **2007,** *450* (7172), 1001-9.

33. Halgren, T. A., Identifying and characterizing binding sites and assessing druggability. *Journal of Chemical Information and Modeling* **2009,** *49* (2), 377-89.

34. Zoete, V.; Meuwly, M.; Karplus, M., Study of the insulin dimerization: binding free energy calculations and per-residue free energy decomposition. *Proteins* **2005,** *61* (1), 79-93.

35. Morris, G. M.; Huey, R.; Lindstrom, W.; Sanner, M. F.; Belew, R. K.; Goodsell, D. S.; Olson, A. J., AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. *Journal of Computational Chemistry* **2009,** *30* (16), 2785-91.

36. Ruah, S. S. H.; Grootenhuis, P. D. J.; Miller, M. T.; McCartney, J.; Van Goor, F.; Numa, M. M. D.; Zhou, J.; Bear, B. Preparation of indolyl cycloalkylcarboxamides as modulators of ATP-binding cassette transporters for use as drugs and biological tools. WO2010054138A2, 2010.

37. Zhang, M.; Huang, Z.; Yu, B.; Ji, H., New homogeneous high-throughput assays for inhibitors of beta-catenin/Tcf protein-protein interactions. *Analytical Biochemistry* **2012,** *424* (1), 57-63.

38. Huang, Z.; Zhang, M.; Burton, S. D.; Katsakhyan, L. N.; Ji, H., Targeting the Tcf4 G13ANDE17 binding site to selectively disrupt beta-catenin/T-cell factor protein-protein interactions. *ACS Chemical Biology* **2014,** *9* (1), 193-201.

39. Park, C. H.; Chang, J. Y.; Hahm, E. R.; Park, S.; Kim, H. K.; Yang, C. H., Quercetin, a potent inhibitor against beta-catenin/Tcf signaling in SW480 colon cancer cells. *Biochemical and Biophysical Research Communications* **2005,** *328* (1), 227-34.

40. Momany, F. A.; Rone, R., Validation of the general purpose QUANTA ®3.2/CHARMm® force field. *Journal of Computational Chemistry* **1992,** *13* (7), 888-900.

41. Nikolovska-Coleska, Z.; Wang, R.; Fang, X.; Pan, H.; Tomita, Y.; Li, P.; Roller, P. P.; Krajewski, K.; Saito, N. G.; Stuckey, J. A.; Wang, S., Development and optimization of a binding assay for the XIAP BIR3 domain using fluorescence polarization. *Analytical Biochemistry* **2004,** *332* (2), 261-73.

CHAPTER 3

MODIFICATION OF INHIBITORS FOR *β*-CATENIN/T-CELL FACTOR PROTEIN-PROTEIN INTERACTIONS

3.1 Introduction

The interfaces between proteins have emerged as a popular therapeutic target. However, finding small-molecule drugs that can disrupt protein-protein interactions is challenging not only because of the flat and featureless surfaces with shallow pockets that are on the protein,[1] but also the complicated thermodynamics involved in the process of protein-protein interactions and ligand binding.[2] Within the past decade, the thermodynamics of molecular binding to guide medicinal chemistry has gained increased attention.[3-7] This chapter focuses on the optimization of the inhibitors of *β*-catenin/T-cell factor 4 (Tcf4) protein-protein interactions (PPIs) using a thermodynamics-guided strategy.

3.1.1 Thermodynamics of Ligand-Protein Interactions

When a drug-like molecule binds to a protein, it becomes less mobile.[8] The resulting loss in entropy, which works against the attractive forces that favor binding, is called the entropic penalty.

The thermodynamic plots of the measured free binding energy (ΔG_{obs}), the measured binding enthalpy (ΔH_{obs}), and the calculated entropic contribution ($-T\Delta S_{obs}$) are

used to demonstrate the driving forces of the process when small-molecule inhibitors bind to a protein. The sum of ΔH_{obs} and $-T\Delta S_{obs}$, ΔG_{obs}, gives a quick assessment of their individual contributions to the binding affinity.

$$\Delta G_{obs} = \Delta H_{obs} - T\Delta S_{obs} \quad (1)$$

Ligand binding enthalpy mostly arises from specific target-ligand interactions such as hydrogen bonds, salt bridges, or directed van-der-Waals contacts, while nonspecific lipophilic interactions cause entropy.[9-12] Typically, a large negative value for ΔH_{obs} is considered to be favorable as it indicates the formation of noncovalent bonds during the binding process. On the other hand, a large negative value for $T\Delta S_{obs}$ (nonspecific hydrophobic effects) is often regarded as disadvantageous. Thus, a balanced thermodynamic profile (larger enthalpic and smaller entropic contributions) has become an important factor in thermodynamics-guided drug design, which often includes the identification of hits having a large enthalpic advantage with a small entropic penalty.

One common effort in entropy management is adding conformational constraints during the synthetic design of the compounds (**Figure 3.1**).[13] As a flexible ligand often exists as multiple conformations, such preorganization attempts to restrain its dynamic motion minimize the unfavorable conformations and hence avoid the unfavorable change in entropy, leading to a more ordered binding configuration.

Within native peptides and peptidomimetics, conformational constraints are widely used to help improve target selectivity,[14-18] favoring the binding to one protein over others that share a high degree of similarity.[2, 19-23] Additionally, the benefits of adding constraints

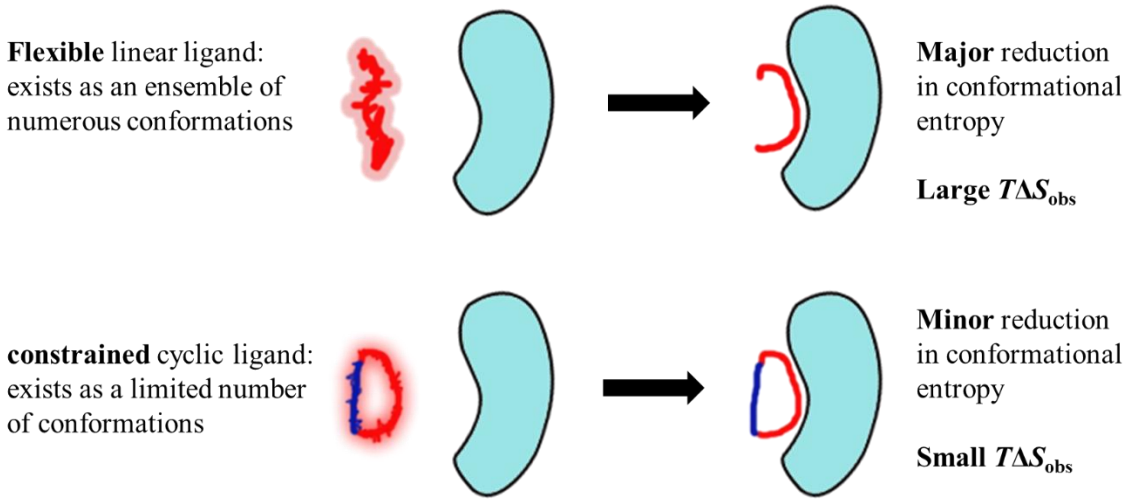


Figure 3.1. Theoretical behavior of conformational entropy in the case of a flexible ligand and corresponding constrained analog. Figure adapted from reference 13.

to peptide and peptidomimetic structures include increased resistance to proteases and *in vivo* metabolic stability,[24-31] along with improved cell membrane permeability.[32-36]

3.1.2 Protein Surface Adaptability

Many protein-protein interaction surfaces are shared with more than one binding partner, and hence the binding surface can be flexible or "plastic".[37] Studies have shown that these interaction surfaces are typically less shaped but adaptable for their binding partners.[38-40] As a result, this adaptability often leads to the opening up of new pockets that could be available for the binding of inhibitors. This effect is consistently observed in the co-crystallized structures of proteins with small molecules,[41, 42] where the small molecule induced new conformations of protein surfaces that were not seen in the native crystal structures of the apoprotein or the protein-protein complex. Inhibitors utilizing this adaptability of protein surface often exhibit a better potency and selectivity. An example is

shown in **Figure 3.2**.[41] However, locating such sites can be difficult. Even if the protein structural information is available, conformational adaptation is required for binding the ligand. Therefore, these "plastic" sites may not be detectable in the unbound receptors.[37]

3.1.3 Previous Work

Previously, peptidomimetic small molecules **3.1** and **3.2** (**Figure 3.3**) were identified as potent and selective inhibitors of the *β*-catenin/Tcf4 PPIs by maintaining the "hot spot interactions" and utilizing the unexploited interactions around the "hot spot". Compound **3.2** has a K_d value of 0.418 μM for binding to *β*-catenin and a K_i value of 1.36 μM to disrupt *β*-catenin/Tcf interactions completely.[44] It also exhibits dual selectivity for *β*-catenin/Tcf over *β*-catenin/cadherin and *β*-catenin/APC interactions.

However, the further optimization of this compound towards a more drug-like molecule was hampered due to several problematic features, mainly coming from the peptide-like structure and the instability of the 2-(1*H*-indol-2-yl)acetate. Therefore, we searched for more potent and more promising drug candidates. Efforts were made to change the inhibitor from a flexible and peptide-like structure to a rigid, more drug-like molecule.

3.2 Results and Discussion

3.2.1 Replacing the Flexible Peptide Backbone with a Rigid Ring Core

Inspection of the docking modes of compound **3.1** gave rise to the proposed structure with a biphenyl core **3.3** shown in **Figure 3.4**. Connectivities of the side chains along with a variety of R groups (halogen, alkyl, and alkoxy substituted phenyl groups)

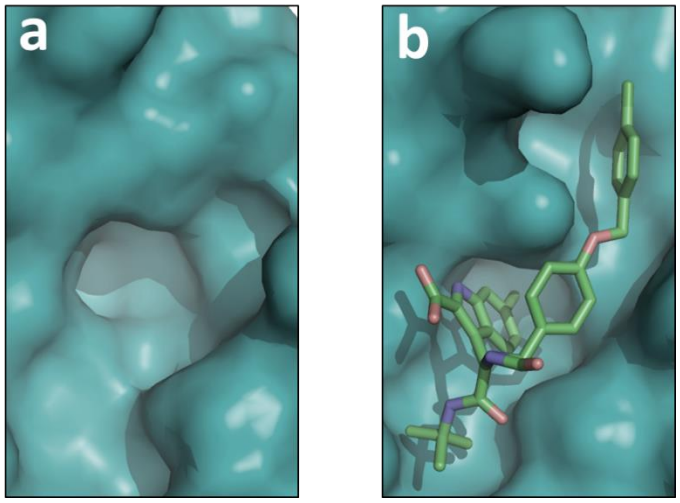
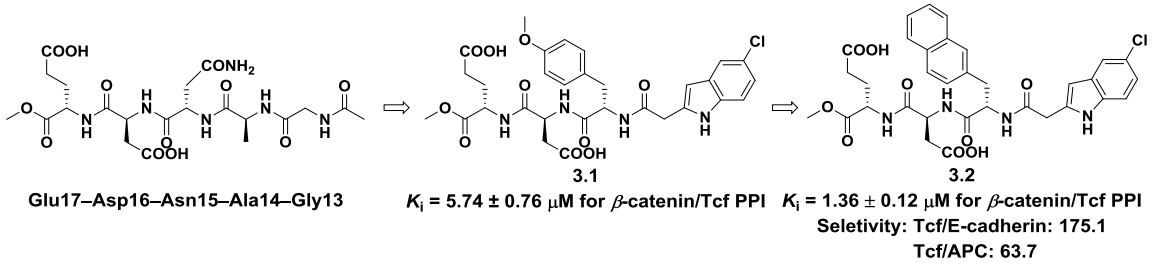


Figure 3.2. An example of surface adaptability.[41] MDM2 (surface is shown in blue) **a**. MDM2 surface with no ligands, PDB id: 3LBK.[43] **b**. MDM2 with YH-300, an MDM2-p53 inhibitor. PDB id: 4MDQ.[41] The small molecule induced a new hydrophobic pocket that was not seen in apoprotein.



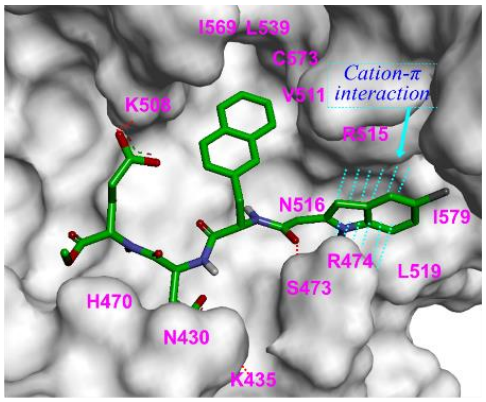


Figure 3.3 Previously identified small-molecule inhibitor for β-catenin/Tcf4 PPIs.

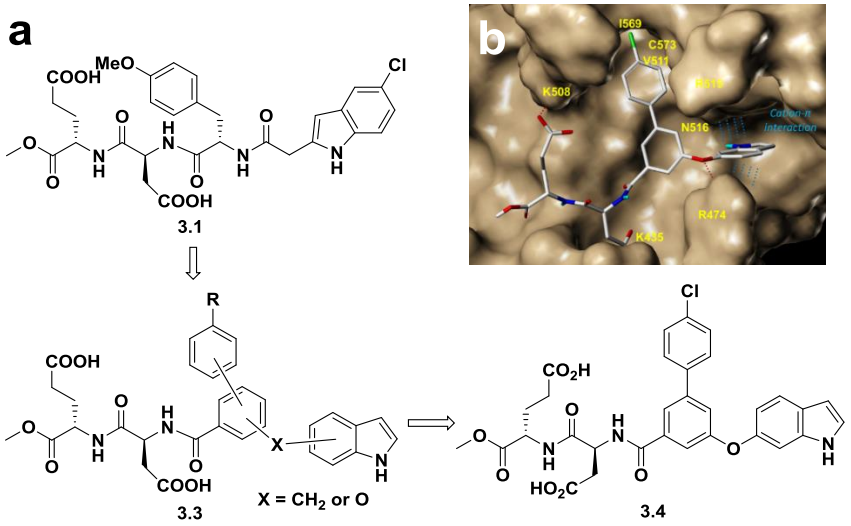


Figure 3.4. **a**. Structural modification from **3.1** to **3.4**. **b**. AutoDock predicted conformation of **3.4**. *β*-Catenin is shown as a space-filled model.

were examined using AutoDock 4.2.[45] Docking results suggested a 1, 3, 5-trisubsituted pattern for the core phenyl ring with an oxygen linker to the 6-position of indole. From the proposed docking mode, this new structure maintains the hydrophobic interactions at pocket B and the indole can still form cation-π interaction with R515 (parallel). Additionally, with the oxygen linker in proximity (2.36 Å to N-H of R474) with A474, a weak H-bond is likely to form and hence adds to the binding affinity.

The synthesis of compound **3.4** is shown in **Figure 3.5**. Using 3-bromo-5-iodobenzoic acid and (4-chlorophenyl) boronic acid as a starting material, this synthesis employed a selective Suzuki coupling followed by an Ullman coupling to assemble the key intermediate **3.9**, which was then coupled with the dipeptide **3.10** reported in the previous chapter.[46-47] Deprotection of the *tert*-butyl esters and Boc group in one pot yielded **3.4**.

Fluorescence polarization (FP) assay was used to determine the potency of compound **3.4** together with other inhibitors (**Table 3.1**).[44, 48-49] To our delight, compound

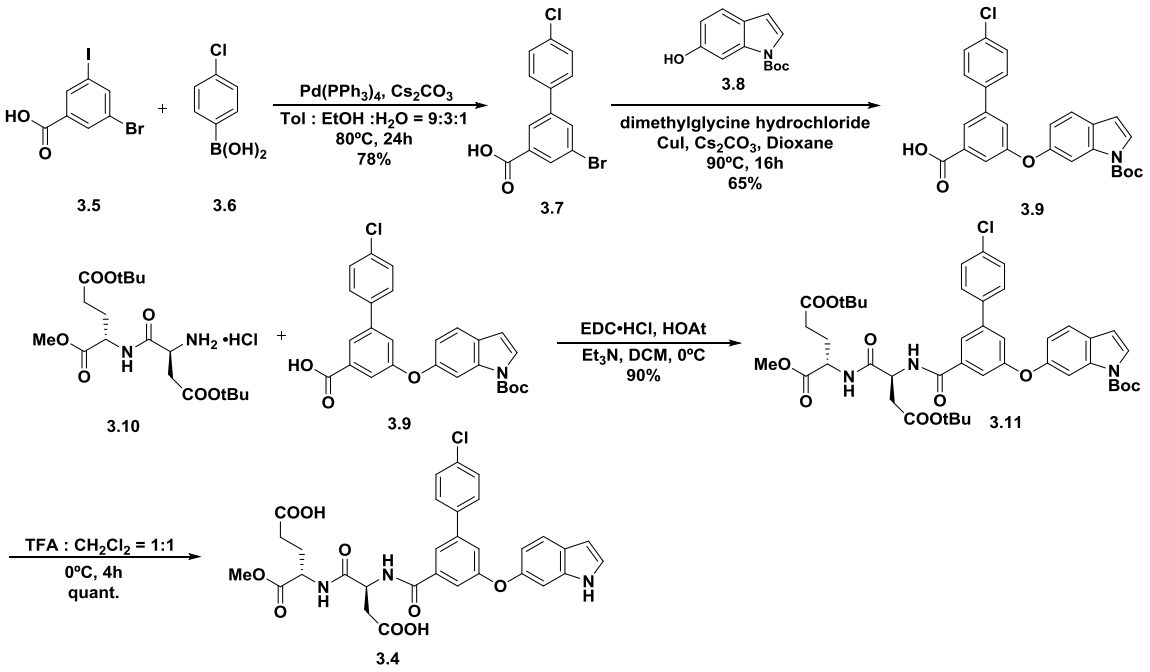


Figure 3.5. Synthesis of compound **3.4**

Table 3.1. The FP competitive inhibition assay results of *β*-catenin/Tcf4 PPIs inhibitors

Entry	$K_i \pm$ SD (μM)
3.1	5.74 ± 0.76[44]
3.2	1.36 ± 0.12[44]
3.4	0.21 ± 0.05
UU-01	3.14 ± 0.48[49]

3.4 exhibited a K_i of 0.2 ± 0.1 μM toward the inhibition of β-catenin/Tcf4 PPIs, which was a 6-fold enhancement to the previously best small molecule inhibitor **3.2**.

3.2.2 Exploration of Arginine Channel (Pocket C)

The synthetic route featuring Ullmann coupling allows the facilitation of different phenols. Therefore it is possible to install various heterocyclic rings onto the scaffold to examine the structural-activity relationship for pocket C. Indazole and benzimidazole were chosen to replace indole for two reasons: 1. they are geometrically similar to indole; 2. they are chemically more stable than indole. Furthermore, pocket C exhibited a hydrophobic area that can potentially be used to increase the binding affinity (**Figure 3.6**).

Following the similar synthesis of **3.4**, a Boc-protected 6-hydroxyl indazole was prepared (**Figure 3.7**). However, the Ullman coupling failed to proceed with the indazole due to the strong electron withdrawing effect of the two nitrogen atoms together with the Boc group. Therefore, benzyl groups were used for both heterocyclic rings. **3.23a-b** were synthesized using a reported method.[50]

Benzyl protected benzimidazoles were synthesized by reducing the nitro group using iron powder in the presence of formic acid followed by ring formation to give **3.26** (**Figure 3.8**).[51] Substituting formic acid with acetic acid in the hope of yielding the 2-methyl benzimidazole resulted in product decomposition upon contact with air, which significantly complicated the purification and lowered the yield. It was then found that protecting the phenol as an acetate could minimize the decomposition and allow the separation and purification of the product. The acetyl group was then hydrolyzed using acidic/basic washes, yielding **3.27**.

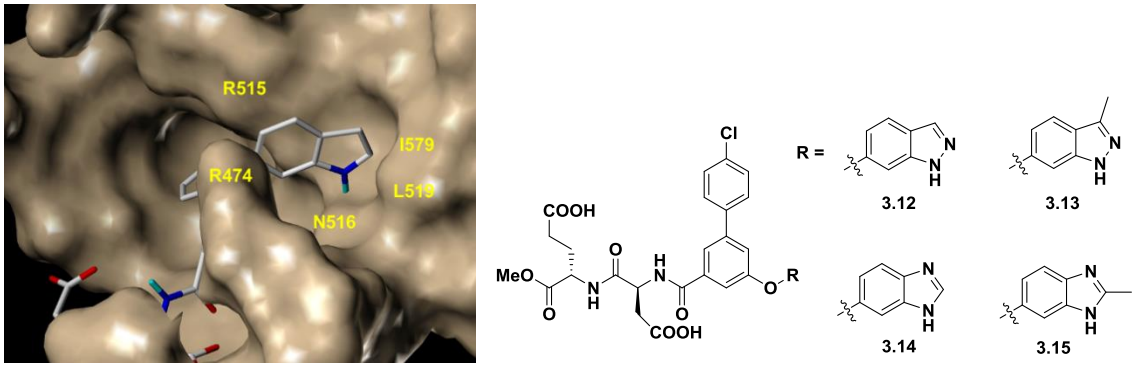


Figure 3.6. Compounds for exploring pocket C.

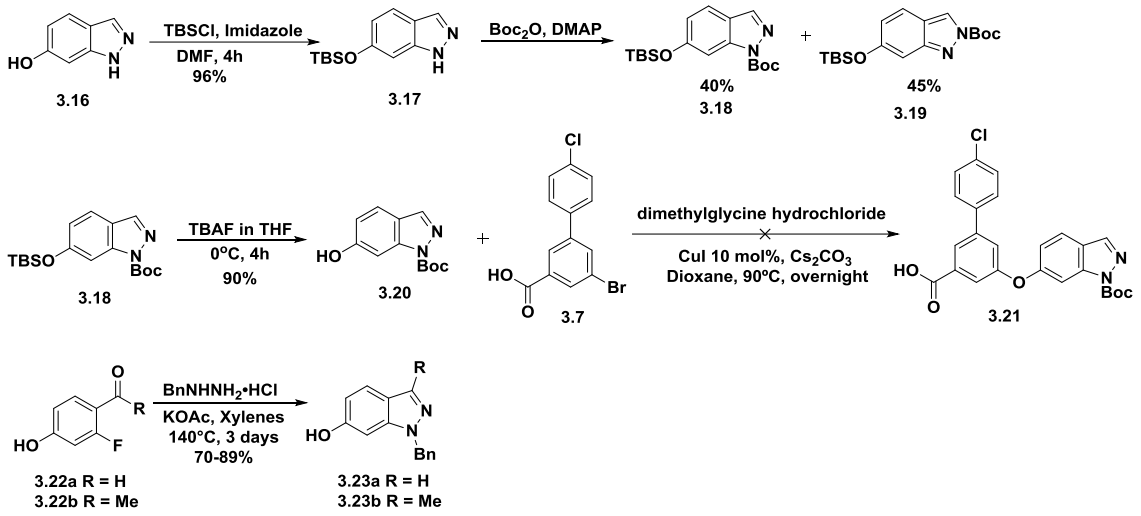


Figure 3.7. Attempts for the synthesis of indazole compounds.

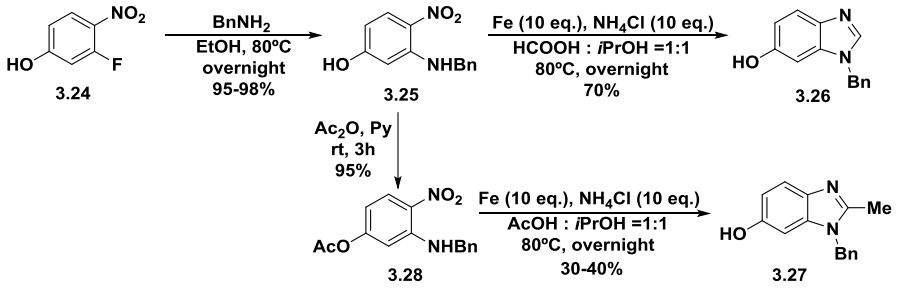


Figure 3.8. Synthesis of benzimidazoles.

Compound **3.26** (1-benzyl-6-hydroxyl benzimidazole) was subjected to Ullman coupling conditions (**Figure 3.9**). The reaction proceeded in moderate yields. After coupling with the dipeptide **3.10**, the benzyl group was found inert to mild hydrogenation. In order to prevent the racemization of the peptides, harsher conditions were not examined. Removal of the *tert*-butyl esters from **3.30** gave **3.31**.

Oxidative debenzylation turned out to be an effective means of removing the benzyl groups on the heterocyclic rings (**Figure 3.10**).[52] Debenzylation on **3.29** was successfully carried out using this condition 85% yield. Interestingly, the reaction rate was found to have a direct dependence on the concentration of potassium *tert*-butyl oxide (*t*BuOK). The reaction proceeded very slowly when the concentration of the base was lower than 0.1 g/mL. Additional base was needed in order to achieve higher conversion.

With the key steps worked out for **3.31** and **3.32**, compounds **3.12-3.15** were synthesized (**Figure 3.11**). The efficiency of the peptide coupling reaction was significantly lowered when protecting groups were absent on the heterocyclic rings.

The AlphaScreen assay was used to evaluate the inhibitory activity of the newly synthesized inhibitors.[48] The results are shown in **Figure 3.12**. In accordance with the AutoDock results, indazole-containing compounds (**3.12** and **3.13**) showed elevated inhibitory activity than **3.4**. Compared to the methyl group on the indazole contributed to a lower K_i presumably from the additional hydrophobic interaction with I579 and L519. On the other hand, benzimidazole adversely affected the K_i of **3.14** and **3.15** due to the hydrophobicity of benzimidazole. A methyl group in the 2 position of the benzimidazole further created an unfavorable steric interaction to fit **3.15** into the pocket. However, the benzyl protected compound **3.31** exhibited an unexpected K_i of 13 nM.

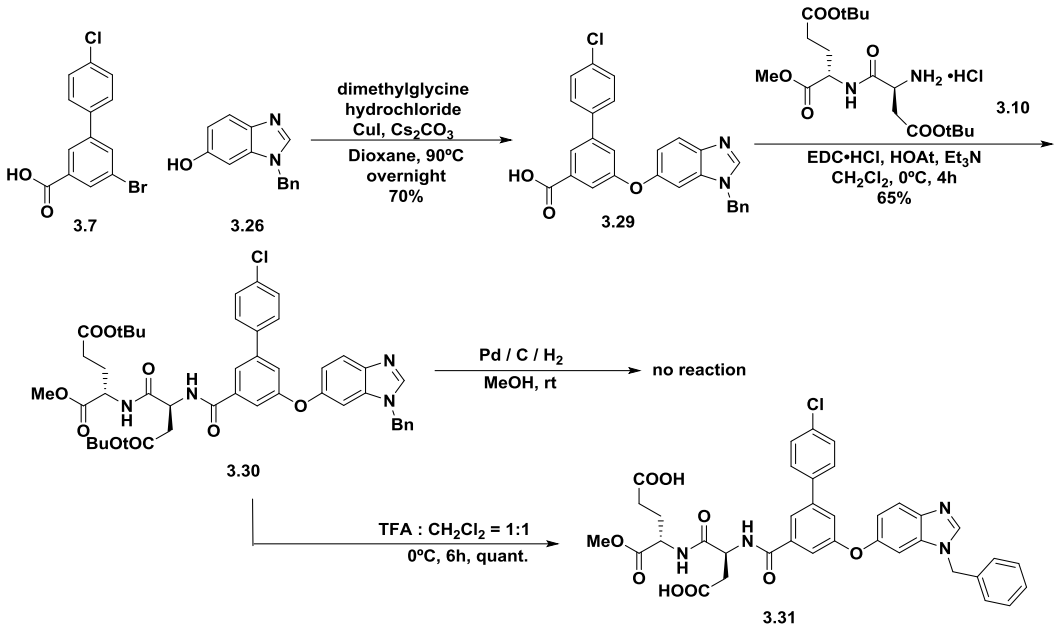


Figure 3.9. Attempted synthesis of benzimidazole compound.

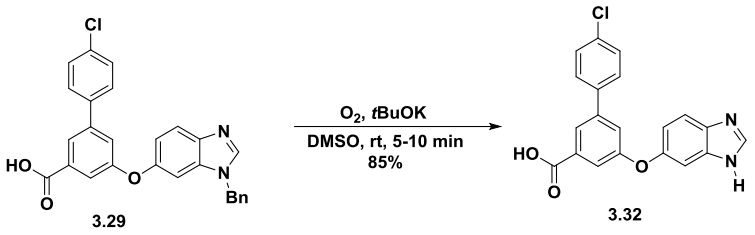


Figure 3.10. Oxidative debenzylation.

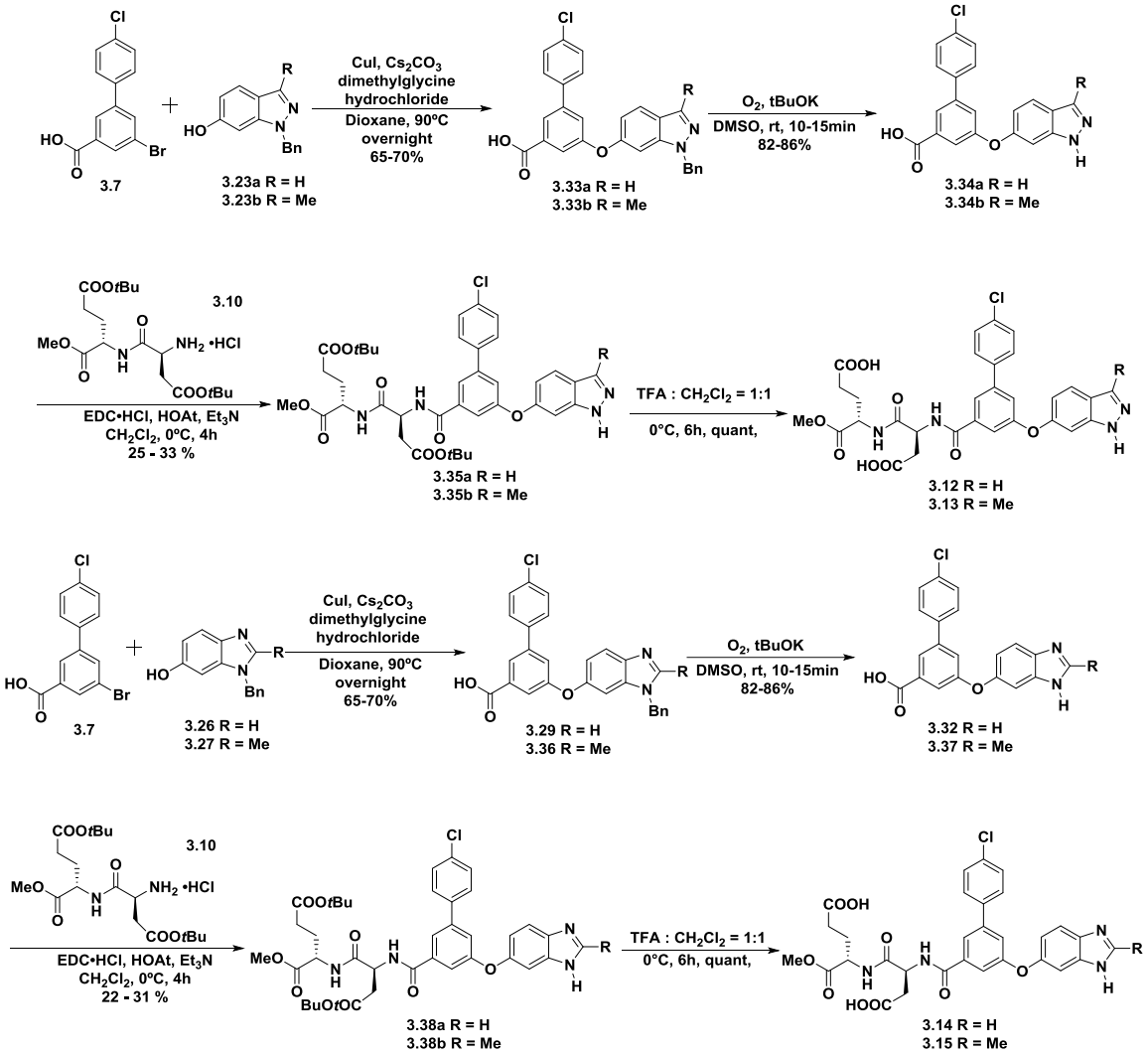
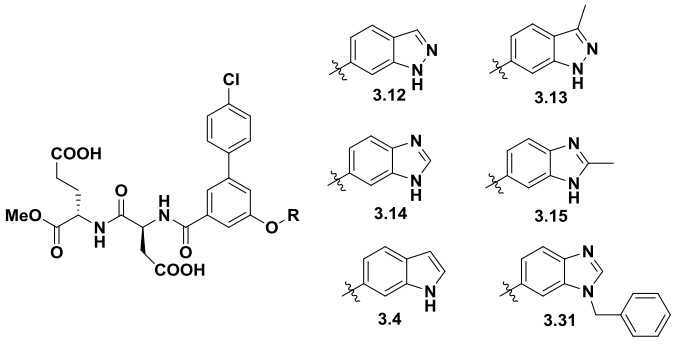


Figure 3.11. Completed synthesis of compounds **3.12-3.15**.



Compound	K_i ± SD (nM)
3.4	200.3 ± 3.4
3.12	96.7 ± 13.9
3.13	77.7 ± 15.5
3.14	436.0 ± 29.3
3.15	572.2 ± 129.6
3.31	13.2 ± 1.7

Figure 3.12. AlphaScreen competitive inhibition assay results.

To investigate the rationale for the exceptional result of **3.31**, a molecular dynamics (MD) simulation study was run by Jonathan Leon Catrow. The R612 adjacent to the arginine channel was discovered to be flexible and uncapped a "hidden" hydrophobic pocket which might contribute to adapting the additional benzyl group from **3.31** (**Figure 3.13**). While this pocket was present to a lesser degree in several crystal structures, MD simulation predicted its accessibility.

Pocket C' is a hydrophobic pocket with H-bond donors consisted of R612, V613, L579 and the backbone of H578 (**Figure 3.14**). In most crystal structures, this pocket is covered with R612 and therefore unavailable. The unveiling of this pocket offers the possibility of utilizing the surface adaptability of *β*-Catenin to design potent inhibitors for *β*-Catenin/Tcf PPIs.

Encouraged by the discovery, a variety of *N*-alkylated benzimidazoles containing compounds were screened by AutoDock 4.2. Some of them were selected to explore the structure-activity relationship (SAR) of the new pocket following the benzimidazole synthesis (**Figure 3.15**). 1-(2-(pyridin-2-yl)ethyl)-1*H*-benzo[*d*]imidazol-6-ol (**3.40f**) was used to synthesize compound **3.43** (**Figure 3.16**) **.**

Unfortunately, this direction was not further pursued due to a major change in the results of the previously tested compounds (**Table 3.2**). The highly concentrated protein stock solution was stored in pH 7.4 buffer, which can cause aggregation of *β*-catenin, leading to the false positive results. Newly expressed proteins (by Dr. Min Zhang) were stored in pH 8.8 buffer, in which the highly concentrated *β*-catenin was stabilized. Additionally, a modification of the assay conditions done by Dr. Min Zhang and Jack Wisniewski provided more consistent performances and more reproducible results.[53]

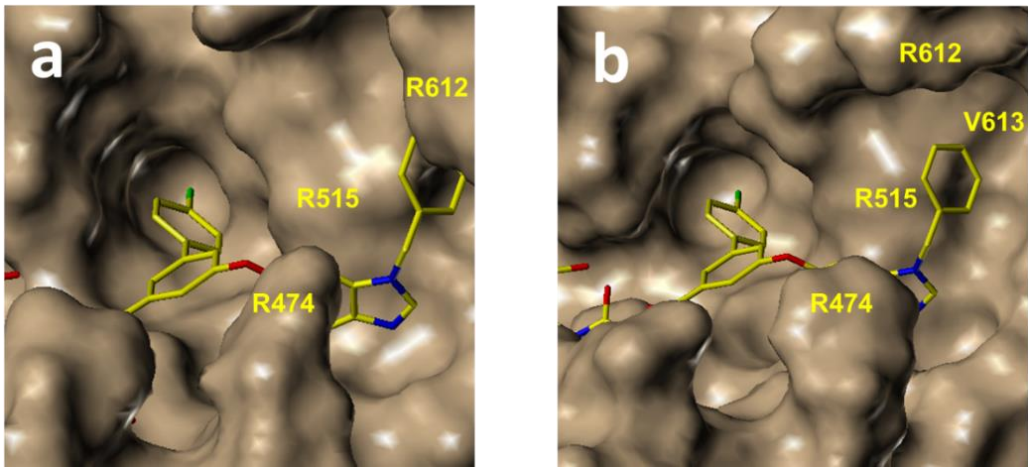


Figure 3.13. *β*-Catenin is shown as a space-filled model. Compound **3.31** is shown in yellow stick model. **a**. **3.13** in the active site of the 2GL7 crystal structure with R612 blocking the benzyl group. **b**. 2GL7 apo simulation cluster 2 from the MD simulation results. R612 moved up opening the pocket C'.

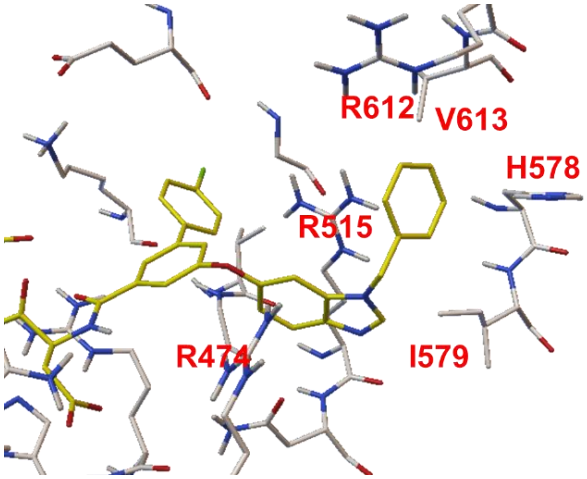


Figure 3.14. Residues of pocket C'. **3.31** is shown in yellow; *β*-Catenin residues are shown in gray.

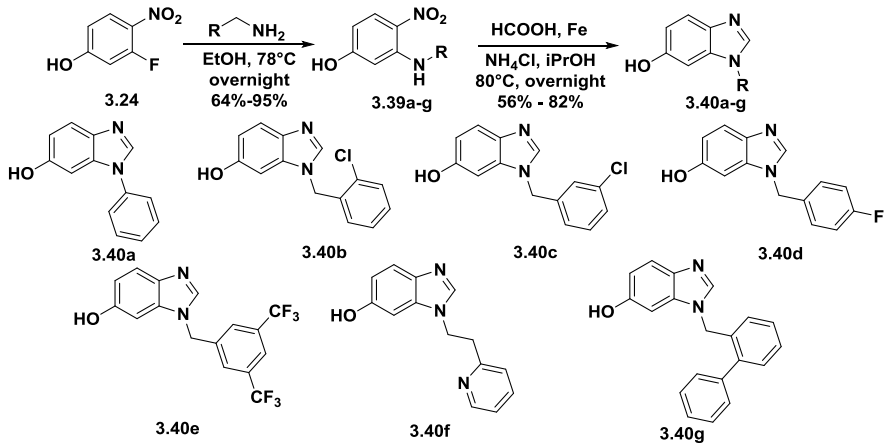


Figure 3.15. Synthesis of *N*-alkylated benzimidazole-6-ols.

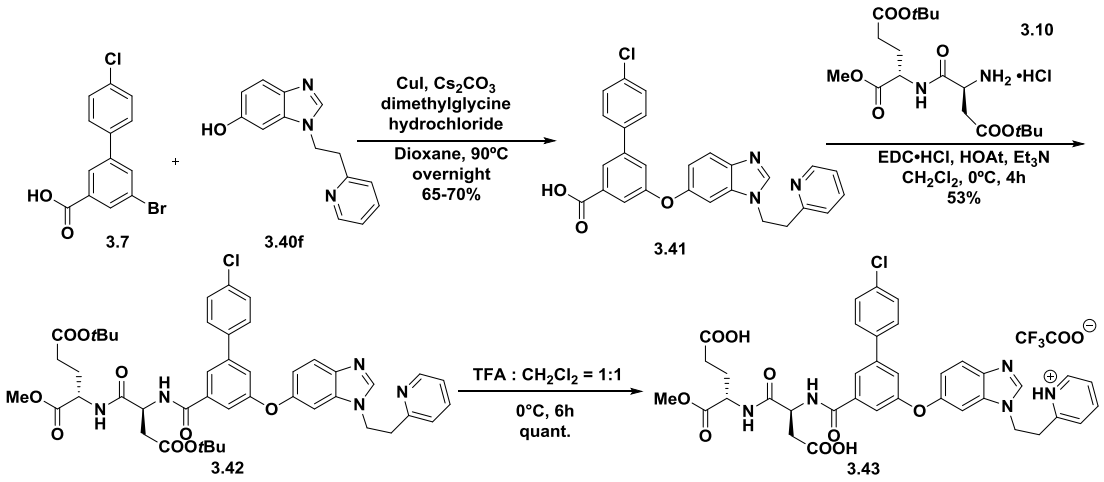


Figure 3.16. Synthesis of compound **3.43**.

Table 3.2 Revised inhibitory activity.

Compound	K_i (μM)		
	AlphaScreen	FP	Originally by AlphaScreen
3.4	0.37 ± 0.10	2.84 ± 0.59	0.20 ± 0.03
3.12	2.03 ± 0.21	5.17 ± 0.84	0.097 ± 0.014
3.13	0.79 ± 0.20	3.04 ± 0.62	0.078 ± 0.016
3.14	16.58 ± 4.93	42.95 ± 1.72	0.436 ± 0.029
3.15	24.67 ± 2.98	45.58 ± 2.12	0.572 ± 0.130
3.31	2.77 ± 0.85	13.28 ± 1.27	0.013 ± 0.002
3.43	24.10 ± 3.24	47.03 ± 2.25	N/A

In the new structure-activity relationship based on the revised inhibitory activities, compound **3.4** remains the most potent inhibitor. The others followed an opposite trend. Indazole containing compounds **3.12** and **3.13** have slightly decreased activity when compared to **3.4**. The 3-methyl group in **3.13** contributed to a better activity. Benzimidazole compounds (**3.14** and **3.15**) exhibit a higher K_i, and the addition of methyl negatively affected the binding affinity. Similar to previous results, with the benzyl group on imidazole, the K_i of **3.31** is lower by 4-fold than that of **3.14**.

3.2.3 Binding Mode Investigation Using Site-Directed Mutagenesis

Site-directed mutagenesis studies were performed by Dr. Min Zhang to determine the importance of each pocket to the binding of the small molecules.

As is shown in **Table 3.3**, none of the compounds responded to any of the mutants. Mutations in the arginine channel had no effect on the inhibitory activity in all cases, and the mutations in the hydrophobic pocket even led to lower K_i's. The results raised the suspicion that the compounds were not in this binding site possibly due to the rigidity of the scaffold.

Table 3.3. Site-directed mutagenesis studies using the FP assay.

	K_i (μM)		
β-catenin	**3.4**	**3.13**	**3.31**
Wild type	2.84 ± 0.59	3.04 ± 0.62	13.28 ± 1.27
V511S	0.62 ± 0.12	1.52 ± 0.21	6.29 ± 0.76
V511S/I569S	1.37 ± 0.19	2.29 ± 0.69	7.22 ± 0.91
R474A	1.61 ± 0.21	4.15 ± 1.09	10.34 ± 0.88
R515A	2.17 ± 0.69	4.50 ± 0.58	10.04 ± 0.90
R474A/R515A	2.18 ± 0.55	5.13 ± 0.86	14.17 ± 1.38

Therefore, compounds **3.44** and **3.45** each with one missing side chain were synthesized in the hope of putting this scaffold back into the pocket and revealing the problematic site of this scaffold (**Figure 3.17**). The inhibitory activities of **3.44** and **3.55** on the wild-type β-catenin/Tcf4 and mutant β-catenin/Tcf4 interactions are shown in **Table 3.4**. As expected, the loss of function has a significant impact on the K_i's of compound **3.44** and **3.45**. However, both compounds displayed better inhibition activities with the mutants, suggesting that this series of molecules completely deviated from the targeted binding site. ITC studies further proved that the compound **3.4** was not binding to the proposed pocket (the K_d of **3.4** binding to wild-type β-catenin is 530 nM, and 590 nM to K435A/K508A mutant.)

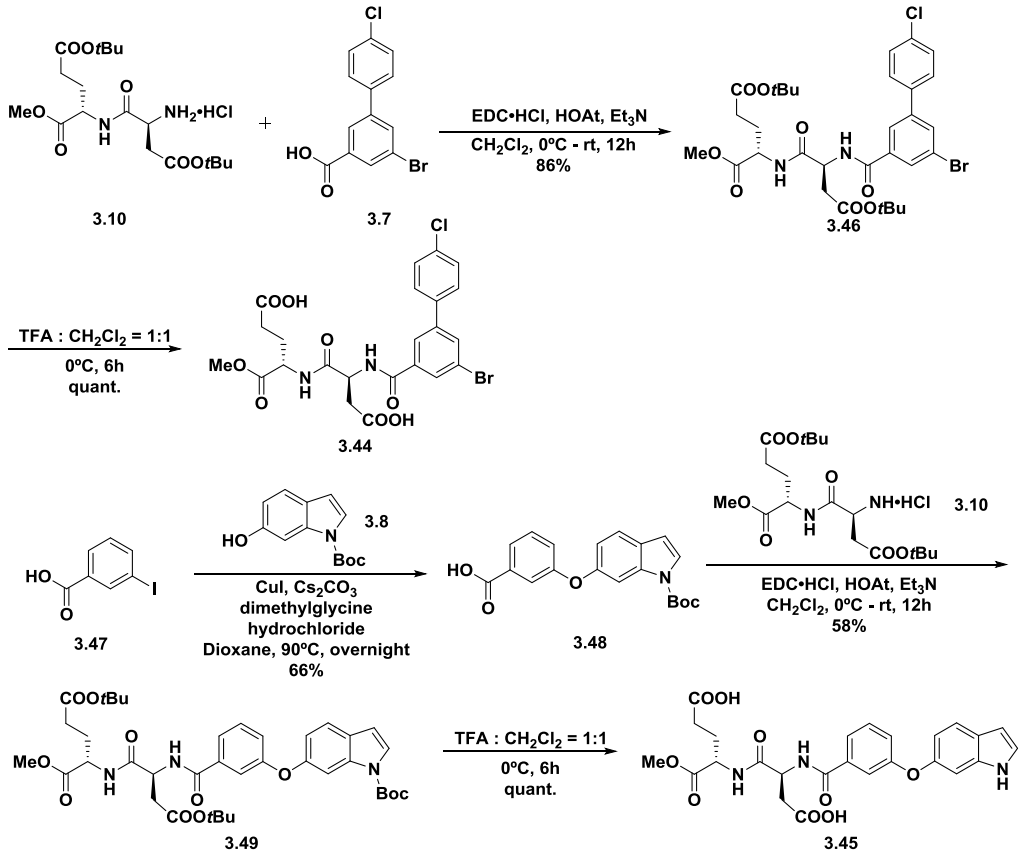


Figure 3.17. Synthesis of **3.44** and **3.45**.

Table 3.4. Site-directed mutagenesis studies and the FP assay to evaluate the binding mode of **3.4**, **3.44** and **3.45**.

	K_i (μM)		
Compound	Wild-type	R474A/R515A	I511S/I569S
3.4	2.84 ± 0.59	2.18 ± 0.55	1.37 ± 0.19
3.44	163.4 ± 6.47	71.75 ± 6.21	58.42 ± 5.33
3.45	19.76 ± 2.78	3.46 ± 0.97	5.00 ± 0.96

3.2.4 Attempts to Validate an Alternative Binding Site

As previously reported, there are three "hot regions" on the *β*-catenin/Tcf4 protein-protein interaction surface.[49] Hot-spot 2 is adjacent to the hot-spot 1 and contains lysine residues (K312 and K345) as key interaction element (**Figure 3.18a**). Besides the similar two residues for charge-charge interactions, the surface area around hot-spot 2 is also equipped with hydrophobic residues I303 and V349, as well as hydrogen bond donor/acceptor T306, W338, and R342 to make it a reasonable site for ligand binding, Therefore, it is likely that the designed inhibitor can bind to the hot-spot 2 in the case of failing to fit in hot-spot 1. Docking studies confirmed that the ligand can bind to the hot-spot 2 area with reasonable conformation (**Figure 3.18b**). The two carboxylic acids can form charge-charge interactions with lysine residues, and the 4-chlorophenyl group stacks with R342 for a cation-π interaction. The indole sits in the hydrophobic pocket formed by I303 and T306, and possibly forming hydrogen bonding interaction with the backbone of Q302.

Four *β*-catenin mutants (K312A, K345A, Y306A and W338A) were made by Dr. Min Zhang. The inhibitory activities of compound **3.13** on Tcf and two *β*-catenin mutants

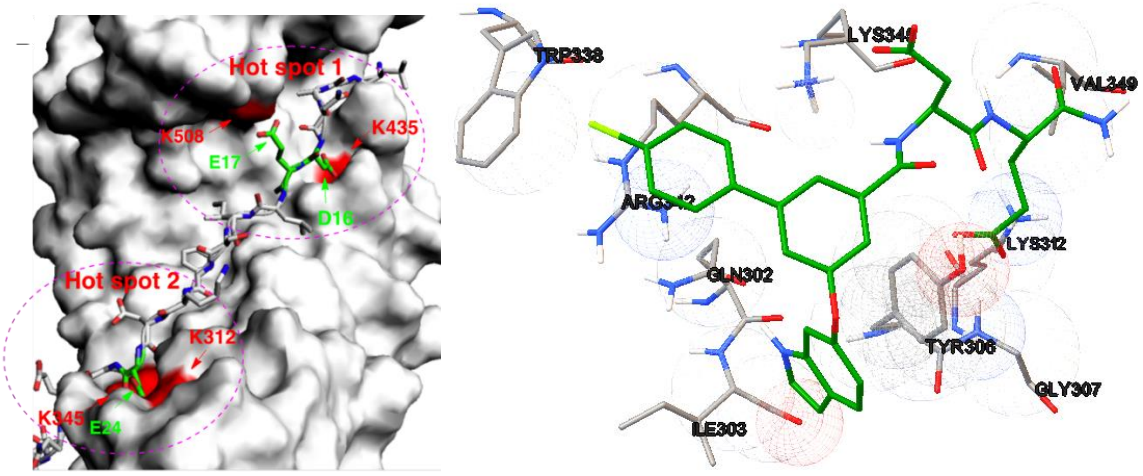


Figure 3.18.[49] Hot spots of the β-catenin/Tcf complex. **a.** "Hot-spots" on the structure of the β-catenin/Tcf complex (PDB id: 2GL7[54]). β-Catenin: surface model (gray), and Tcf: stick model. "Hot-spot" 1: K435 and K508 of β-catenin. "Hot-spot" 2: K312 and K345 of β-catenin. **b.** Proposed docking mode for ligand at the active site from AutoDock 4.2. Ligand **3.4** is shown in green, and the active site residues are shown in gray.

(Y306A and W338A) interactions were tested (**Table 3.5**). The K_i's for the mutant proteins are indeed higher than the wild-type, indicating a decrease in the binding affinity after the lost-of-function of the key amino acid side chains. Although a 4-fold difference is hardly significant, this preliminary study suggested that the inhibitors are very likely to bind to this hot-spot 2 area for its more open pockets as compared to the deep and narrow ones at hot-spot 1.

More studies need to be performed for the further validation of the binding mode. Moreover, the K312A mutation causes a substantial drop of the binding affinity between β-catenin and Tcf4. Because of this decrease in the association constant (K_A), it makes the FP assay no longer suitable for evaluating the inhibitory activity of the ligands. Therefore, ITC experiments are recommended for the validation of the hot-spot 2 binding.

Table 3.5. The FP assay for *β*-catenin/TCF interaction.

	K_i (μM)		
Compound	Wild-type	Y306A	W338A
3.13	3.04 ± 0.62	12.39 ± 2.77	11.61 ± 3.21

3.2.5 Inhibitor Selectivity Studies

The selectivity assay was established in the Ji laboratory using FP to quantify inhibitor selectivity between *β*-catenin/Tcf, *β*-catenin/E-cadherin, and *β*-catenin/APC interactions.[55] The results are shown in **Table 3.6**. The selectivities of **3.4** for *β*-catenin/Tcf over *β*-catenin/E-cadherin interactions and *β*-catenin/APC interactions are 10- and 16-fold, respectively. Comparing to the previously reported best inhibitor **3.2**,[44] compound **3.4** shows significantly decreased selectivities, which on the other hand proves that the inhibitors targeting the hot-spot 1 region can indeed achieve higher selectivities.

3.2.6 Rationale of the Off-Target Outcome

To investigate the cause of the off-target binding of the newly designed inhibitors, the docking modes of the previously designed (**3.2**) and newly designed (**3.4**) inhibitors were extracted and overlapped shown in **Figure 3.19**.

Prominently, compound **3.4** reaches deeper in pocket B than compound **3.2** but comes short in pocket C. The rigidity of compound **3.4** does not allow the free adjustment of side chains. Therefore, the hydrophobic biphenyl core has to stay elevated from the protein surface in order to allow the side-chains to fit into the pockets and hence be exposed to the aqueous solution. Taking off either side chain does not increase the likelihood of placing the core or the other side-chain in a more favorable position due to the lack of

Table 3.6. The FP selectivity assay to determine inhibitor selectivity of the *β*-catenin/Tcf4 inhibitors.

Compounds	$K_i \pm$SD (μM)			Selectivity	
	β-catenin/ Tcf4	*β*-catenin/ E-cadherin	*β*-catenin/ APC	Tcf/ E-cadherin	Tcf/APC
3.4	2.84 ± 0.59	30.69 ± 2.62	46.81 ± 3.22	10.8	16.5
3.2	1.36 ± 0.12	238.1 ± 2.3	86.63 ± 0.64	175.1	63.7

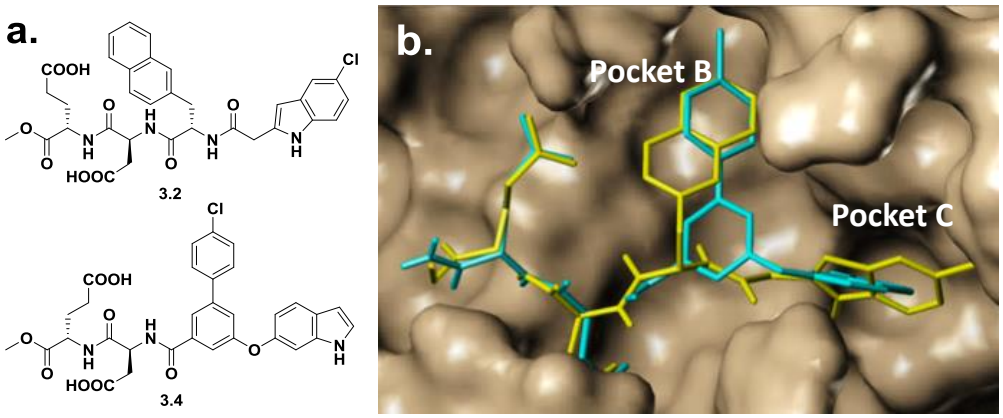


Figure 3.19. Comparison of the two scaffolds. **a.** Structures of the previously designed inhibitor **3.2** and the inhibitor **3.4** designed in this chapter. **b.** Overlapped docking mode of **3.2** (shown in yellow) and **3.4** (shown in cyan). *β*-Catenin is shown as a space-filled model.

rotatable bonds. On the other hand, pocket B is a narrow hydrophobic area. Attempting to insert additional hydrophobic groups into the narrow pocket can possibly cause increased entropy (protein structure change), which disfavors the thermodynamics of the binding.

3.3 Conclusion

Proteins have large surfaces and interact dynamically with a complicated network of other partners. Hence the discovery of small-molecule inhibitors of PPIs is challenging. In this chapter, we attempted to use the thermodynamics of protein-ligand interactions as a guideline to modify the known inhibitors. A new series of inhibitors with rigid core structures were synthesized (**Figure 3.20**). Unfortunately, the overly rigid structure was unable to fit into the target binding site. Preliminary results suggest binding to hot-spot 2. The binding of the newly designed inhibitors relies largely on the heterocyclic ring and leads to a disruption of β-catenin/Tcf interaction with low micromolar K_i's. The selectivity of compound **3.4** was lower than that of **3.2** due to the fact that hot-region 2 is not selective.

3.4 Future Directions

3.4.1 Synthesis of New Compound with a Pyridine Core

The inhibitors designed in this chapter resulted in the disruption of β-catenin/Tcf4 interactions. This unexpected result provided a starting point to for the exploration of the hot-spot 2 binding site. Compared to the deep cleft present in hot-spot 1, hot-spot 2 is flatter in shape and solvent exposed. Therefore, replacing the benzene core with a less hydrophobic structure will stabilize the inhibitor on the surface of protein. Also, the hydrogen bond acceptor site from pyridine can potentially form interactions with protein

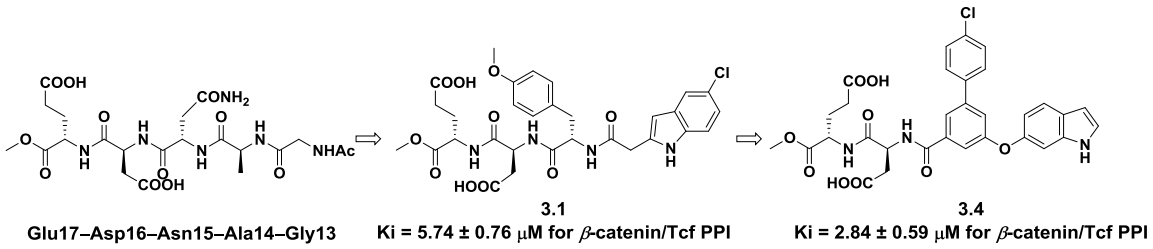


Figure 3.20. From peptide to small molecules.

side chains and/or backbone. On the choice of heterocyclic rings, the 1-benzyl protected 3-methyl indazole derivative is given a high hope for this new compound because of its enhanced chemical stability compared to indole. Additionally, the benzyl group will add to the inhibitory activity according to the SAR of the benzimidazoles. The synthesis of the new pyridine containing compound is shown in **Figure 3.21**.

The key intermediate **3.55** was synthesized through a stepwise functionalization of 4-chloropyridine (**3.50**) to install the cyano group and chloride with moderate yields. The indazole was installed by Ullman coupling reaction, and hydrolysis of the cyano group yielded the carboxylic acid **3.57**, which was then coupled with the dipeptide **3.10**. The *tert*-butyl esters were deprotected in trifluoroacetic acid to afford compound **3.59**. To date, the inhibitory activity of this compound has not been investigated.

3.4.2 Benzodiazepines as New Core Structures

As previously discussed, we believed that the overly rigid structure disfavored the placement of the side chains in the active site. Thus, an ideal new core structure should possess both rigidity as well as a certain degree of flexibility.

Based on the docking mode of **3.2**, the 7-membered diazepine in **3.60** is proposed to meet those requirements (**Figure 3.22**). The 7-member ring keeps both side chains in

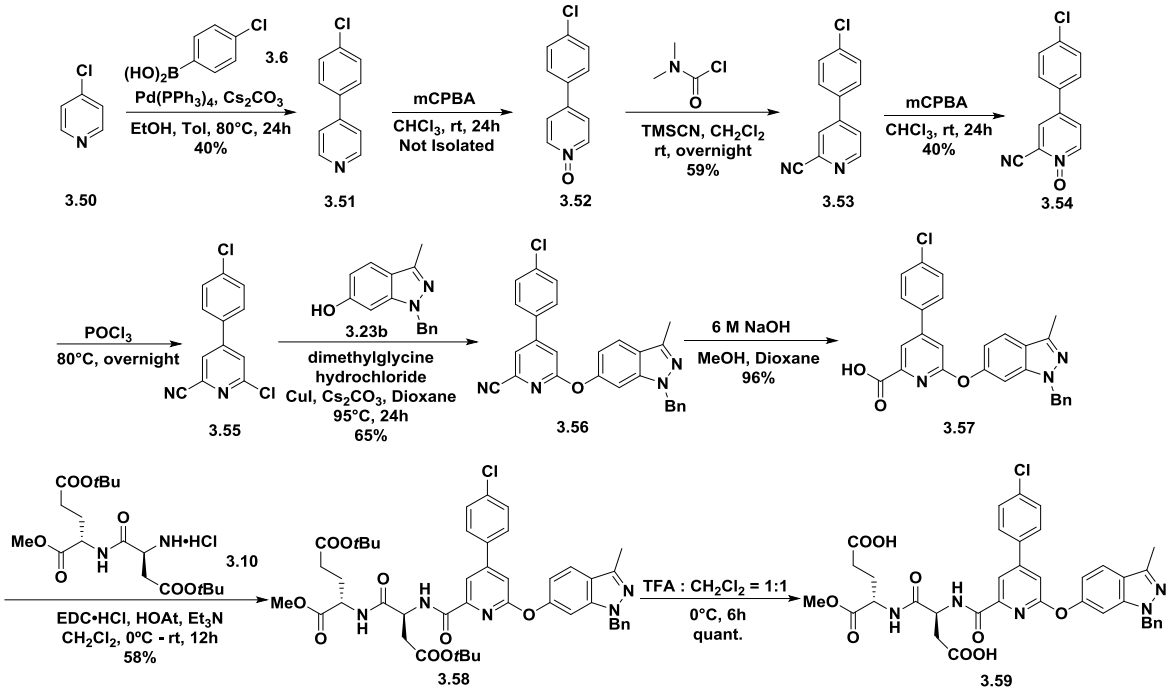


Figure 3.21. Synthesis of the inhibitor with pyridine core.

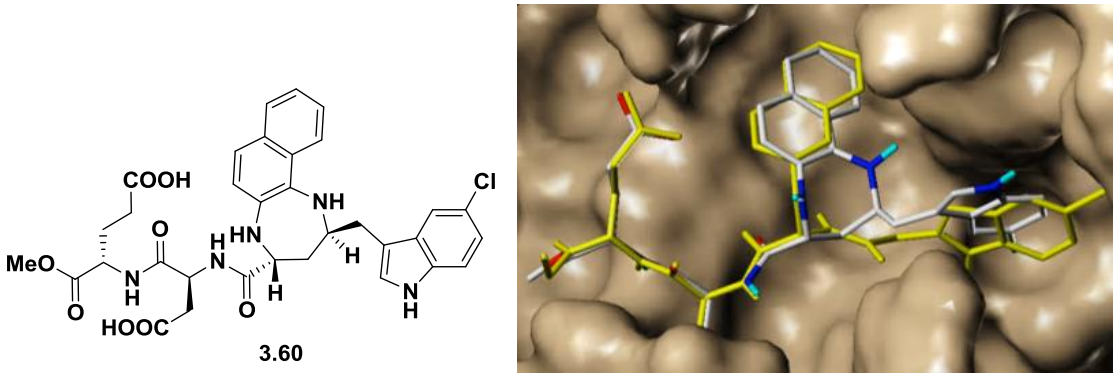


Figure 3.22. Structure and overlapped docking mode of the diazepine containing compound **3.60** (shown in gray) and **3.2** (shown in yellow). *β*-Catenin is shown as a space-filled model.

close proximity to their original position after minimization and possesses both the degree of rigidity and the flexibility to adapt to the active site.

Other forms of diazepine-containing structures were also screened, and compound **3.61** was chosen as a potential starting molecule (**Figure 3.23)**. The 7-member ring is also capable of mimicking the peptide conformation, and the amide groups add to the hydrophobicity of the compound, facilitating the surface binding.

A proposed synthesis is shown in **Figure 3.24**. The Ugi reaction would be used to assemble the precursor **3.67** as a racemic mixture for the diazepine. The removal of the Boc group would trigger the ring-closing in one pot. Following peptide coupling to give **3.70**, the racemic mixture would be converted to a mixture of diastereomers.

3.5 Methods and Experiments

3.5.1 Ligand Docking Using AutoDock 4.2[45]

In AutoDock studies, only the polar hydrogen atoms remained on the protein structure, and Kollman united atom charges were assigned. The 3D structures of the ligands were built, and the partial atomic charges were calculated using the Gasteiger–Marsili method. The rotatable bonds in the ligands were defined using AutoTors, which also united the nonpolar hydrogens and partial atomic charges to the bonded carbon atoms. The grid maps were calculated using AutoGrid. The AutoDock area was defined to include all the residues of the Tcf4 $G^{13}ANDE^{17}$ binding site, and the grid spacing was set to 0.375 Å. Docking was performed using the Lamarckian genetic algorithm, and the pseudo-Solis and Wets method were applied to the local search. Each docking experiment was performed 100 times, yielding 100 docked conformations. Other settings were the standard default

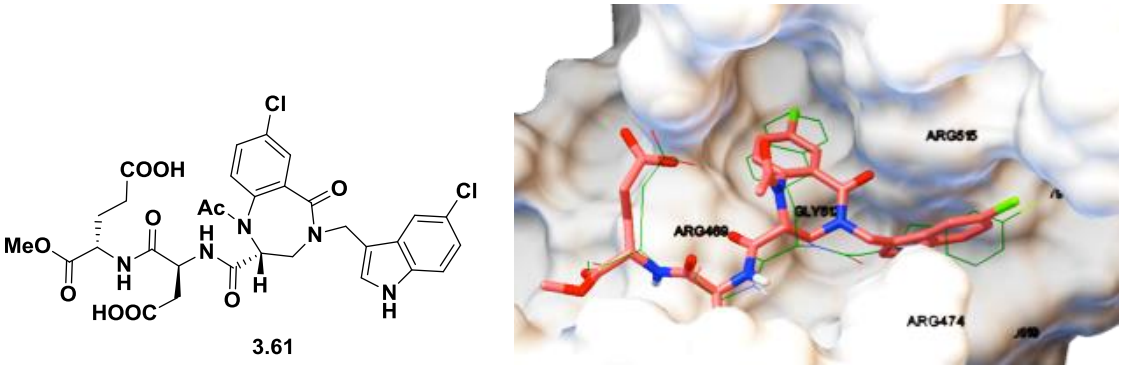


Figure 3.23. Structure and the docking mode of compound **3.61** (shown in red) overlapped with **3.2** (shown in green).

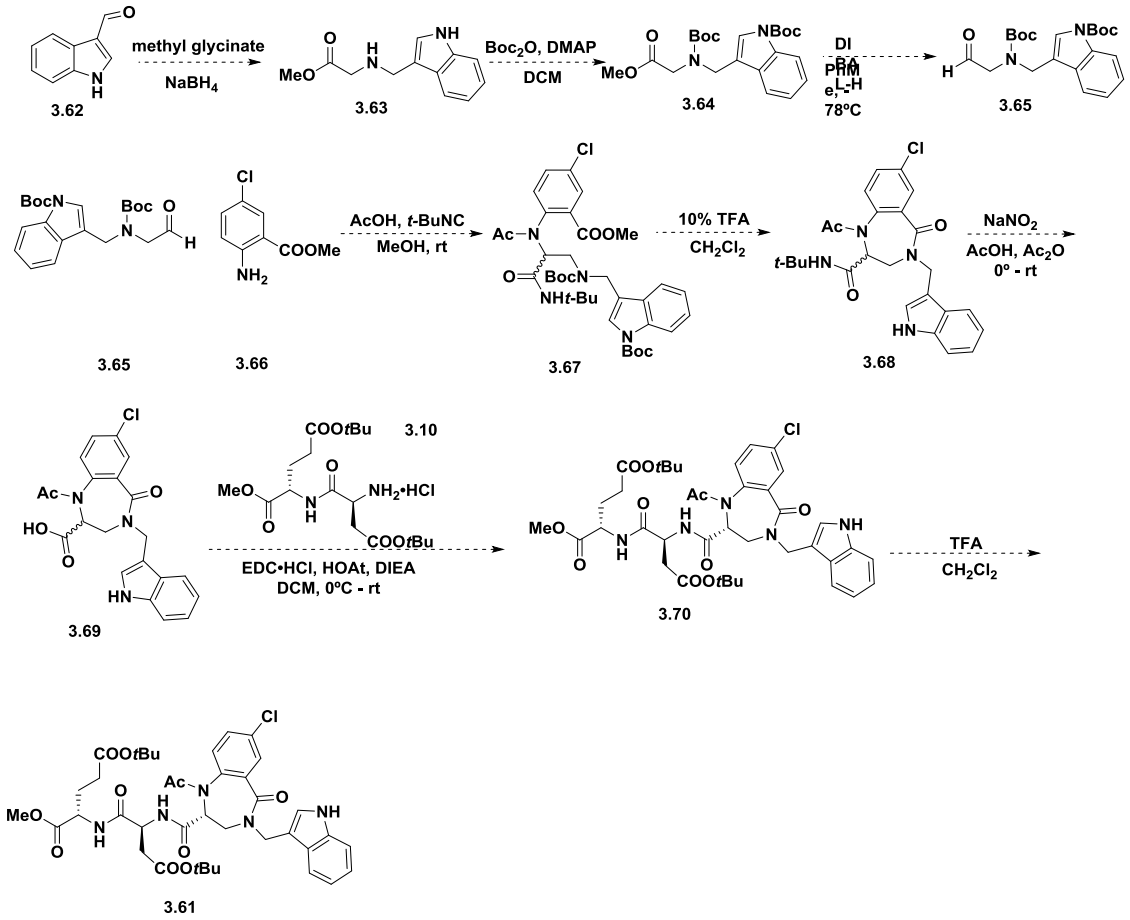


Figure 3.24. Proposed synthesis of diazepine containing compounds.

parameters. All of the ligands followed the same docking protocol. The results of the docking experiments were evaluated by the auxiliary clustering analysis and a visual inspection to match the proposed binding elements.

3.5.2 FP Competitive Inhibition Assay[48, 55]

Experiments were performed in 96-well Microfluor 2 black plates on a Synergy 2 plate reader (Biotek). The polarization was measured at room temperature with an excitation wavelength at 485 nm and an emission wavelength at 535 nm. The FP experiments were performed in an assay buffer of 137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 2 mM KH_2PO_4, 100 μg/mL of bovine gamma globulin, and 0.01% Triton-X 100. The final reaction volume was set to 100 μL. For the β-catenin/Tcf assay, 10 nM of β-catenin (residues 142–686) was incubated with 2.5 nM of *C*-terminally fluorescein-labeled human Tcf4 (residues 7–51) for 30 min at 4 ℃, and then different concentrations of the tested peptides and compounds in assay buffer were added. The negative control (equiv to 0% inhibition) refers to 2.5 nM Tcf4 fluorescence tracer and 10 nM β-catenin in assay buffer without the tested compound. The positive control (equiv to 100% inhibition) refers to only 2.5 nM Tcf4 fluorescence tracer in assay buffer. For the β-catenin/cadherin assay, 150 nM human β-catenin (residues 138−686) was incubated with 5 nM of *C*-terminally fluorescent-labeled human E-cadherin (residues 819−873) in assay buffer for 30 min at 4 ℃. The negative control refers to 5 nM E-cadherin fluorescence tracer and 150 nM β-catenin in assay buffer with no inhibitor presenting. The positive control refers to 5 nM E-cadherin fluorescence tracer in assay buffer. For the β-catenin/APC-R3 assay, 2000 nM human β-catenin (residues 138−686) was incubated with 5 nM of *C*-terminally

fluorescent-labeled human APC-R3 (residues 1477−1519) in assay buffer for 30 min at 4 ℃. The negative control refers to 5 nM APC-R3 fluorescence tracer and 2000 nM *β*-catenin in assay buffer without the tested compound. The positive control refers to 5 nM APC-R3 fluorescence tracer in assay buffer. Each assay plate was covered black and gently mixed on an orbital shaker at 4 ℃ for 2.5 h to reach equilibrium before the polarization values were read. The background of the tested inhibitors was corrected by subtracting the raw intensity values of the sample background well (all components except probe) from the raw intensity values of the corresponding test wells (all components). The IC_{50} values were determined by GraphPad Prism 5.0. The K_i values were derived from the IC_{50} values.[55] All of the experiments were performed in triplicate and carried out in the presence of 1% DMSO for small-molecule inhibitors. The results were expressed as mean ± standard deviation. The Tcf/cadherin selectivity ratios were calculated based on the respective K_i value of *β*-catenin/E-cadherin interactions over that of *β*-catenin/Tcf4 interactions. The Tcf/APC selectivity ratios were calculated based on the respective K_i value of *β*-catenin/APC-R3 interactions over that of *β*-catenin/Tcf4 interactions.

3.5.3 AlphaScreen Assay[48]

Experiments were performed in white opaque 384-well plates from PerkinElmer (Waltham, MA), and the samples were read on a Synergy 2 plate reader (Biotek, Winooski, VT) with a sensitivity setting of 200 using AlphaScreen protocol with excitation at 680 nm and emission at 570 nm. All dilutions were made in 1 × assay buffer containing 25 mM HEPES (pH 7.4), 100 mM NaCl and 0.1% BSA to minimize nonspecific interactions. In the AlphaScreen competitive inhibition assay, 20 nM of *N*-terminal His_6-tagged *β*-catenin

(residues 142-686) was incubated with 5 nM of *C*-terminal biotinylated Tcf4 45-mer (residues 7-51) for 30 min at 4 ℃, and then different concentrations of the tested inhibitors in the assay buffer were added to make a final volume of 20 μL. The mixture was incubated at 4 ℃ for 2 h. Donor and acceptor beads were added to a final concentration of 10 μg/mL in 25 μL of assay buffer. The mixture was incubated at 4 ℃ for 1 h. The IC_{50} value was determined by nonlinear least-square analysis of GraphPad Prism 5.0. The K_i values were derived from the IC_{50} values by a reported method.[56] Experiments were performed in triplicate and carried out in the presence of 1 % DMSO.

3.5.4 ITC Experiments

ITC measurements were performed at 30 ℃ using a VP-ITC (Microcal, GE Healthcare Life Sciences). The concentrations of inhibitor and β-catenin were set to 80 μM and 8 μM, respectively, in a buffer of 20 mM HEPES, pH 7.5, 200 mM NaCl, 1mM TCEP, 0.01% Triton X-100, and 8% glycerol. Each titration experiment was initiated by a 2 μL injection, followed by 30–35 injections with 8 μL per injection. Blank titrations, which were carried out by injecting the compound into the buffer, were subtracted from each data set. The association constant K_A, enthalpy change (ΔH), and stoichiometry N were obtained from fitting the data using the Origin software package (Microcal). The K_d, the free energy change ΔG, and the entropy change ΔS were obtained from the basic thermodynamic relationships $K_d = K_A^{-1}$, $\Delta G = -RT\ln K_A$, and $\Delta G = \Delta H - T\Delta S$.

3.5.5 Site-Directed Mutagenesis Studies

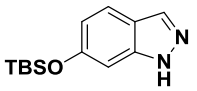
β-Catenin mutants V511S, V511S/I569S, R474A, R515A, R474A/R515A, K312A, K345A, T306A and W338A were generated using the overlapping PCR technique. The template for the mutagenesis reactions was the full-length wild-type *β*-catenin in pET-28a. KOD Hot start DNA polymerase (Novagen) was used for all the experiments. Mutants were confirmed by direct sequencing (Core Facility, University of Utah). The templates used for *β*-catenin double mutation were V511S for V511S/I569S and R474A for R474A/R515A, respectively. The double mutants were again confirmed by direct sequencing. Following the confirmation of the sequence, *β*-catenin mutants were cloned into a pET-28 vector and transformed into *E. coli* BL21 DE3.

3.5.6 Chemistry

All experiments were conducted under anhydrous conditions in an atmosphere of argon, using flame-dried apparatus and employing standard techniques in handling air-sensitive materials. Dichloromethane (CH_2Cl_2), acetonitrile (CH_3CN), tetrahydrofuran (THF), dimethylformamide (DMF) were degassed with nitrogen and passed through JC Meyer solvent systems. All reagents were purchased from Sigma-Aldrich and used as received. Aqueous solutions of sodium bicarbonate ($NaHCO_3$), sodium chloride (brine), and ammonium chloride (NH_4Cl) were saturated. Analytical thin layer chromatography was visualized by ultraviolet light. Flash chromatography was performed on SilicaFlash$^{@}$ F60 silica gel (230–400 mesh). ^{1}H NMR spectra were recorded using a Varian Unity Inova 500 (500 MHz) or a Varian Unity Inova 300 (300 MHz). ^{13}C NMR spectra were recorded using a Varian Unity Inova 500 (125 MHz) or Varian Unity Inova 300 (75 MHz). The ^{1}H

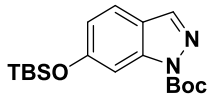
and ^{13}C NMR spectra were referenced to the residual solvent signals (7.26 ppm for ^{1}H and 77.0 ppm for ^{13}C in $CDCl_3$; 2.50 ppm for ^{1}H and 39.5 ppm for ^{13}C in DMSO-d6). All melting points were determined on a Büchi M560 melting point apparatus and uncorrected. Low (MS) and high (HRMS) mass spectra were determined on a Micromass Quattro II (ESI/APCI-TOF) at the University of Utah, Department of Chemistry Mass Spectrometry Facility. Optical rotation of final compounds was determined on a Perkin-Elmer PE-343 polarimeter.

3.5.7 Synthetic Routes, Procedures and Characterizations



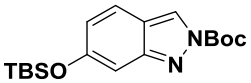
3.17

6-((*tert*-Butyldimethylsilyl)oxy)-1*H*-indazole (3.17). To a stirred solution of 5-hydroxyl-1*H*-indazole (0.5 g, 3.727 mmol) in 10 mL DMF under room temperature was added TBSCl (0.674 g, 4.472 mmol) and imidazole (0.38g, 5.59 mmol). The reaction was left to react overnight. Upon completion, the reaction mixture was poured into 30 mL 1 M HCl, and extracted with ethyl acetate (40 mL). The organic layer was washed with 2 M HCl (aq) (30 mL), saturated $NaHCO_3$ (aq) (30 mL), brine (2×30 mL) and then dried over Na_2SO_4. The solvent was evaporated under reduced pressure to afford 0.89 g product as colorless oil (96%). ^{1}H NMR (500 MHz, Chloroform-*d*) δ 7.99 (s, 1H), 7.59 (dd, *J* = 8.6, 0.7 Hz, 1H), 6.92 – 6.86 (m, 1H), 6.75 (dd, *J* = 8.7, 2.0 Hz, 1H), 1.01 (s, 9H), 0.23 (s, 6H). ^{13}C NMR (126 MHz, $CDCl_3$) δ 155.7, 141.5, 134.8, 121.6, 118.7, 117.1, 99.2, 25.9, 18.5, -4.2. MS (ESI) *m/z* = 249.4 $[M+H]^+$.



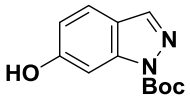
3.18

***tert*-Butyl 6-((*tert*-butyldimethylsilyl)oxy)-1*H*-indazole-1-carboxylate (3.18).** To a stirred solution of 6-((*tert*-butyldimethylsilyl)oxy)-1*H*-indazole (**3.17) (**0.265 g, 1.07 mmol) in 15 mL dichloromethane at room temperature was added Boc_2O (0.26 g, 1.2 mmol) and N, N-dimethyl-4-aminopyridine (10 mg, 0.08 mmol). The reaction was completed in 5 min. The reaction mixture was poured to 2 M HCl (aq) (30 mL) and extracted with dichloromethane (2 × 20 mL). The organic layers were combined and concentrated under reduced pressure. The residue was purified by column chromatography (hexanes : ethyl acetate =10:1) to yield 0.15g **3.18** (40%) and 0.16g **3.19** (45%). ^{1}H NMR (500 MHz, Chloroform-*d*) δ 8.50 (d, *J* = 1.1 Hz, 1H), 7.47 (dd, *J* = 9.2, 0.8 Hz, 1H), 6.98 (dt, *J* = 1.9, 0.9 Hz, 1H), 6.72 (dd, *J* = 9.1, 2.0 Hz, 1H), 1.69 (s, 9H), 1.51 (s, 3H), 0.98 (s, 9H), 0.23 (s, 6H). ^{13}C NMR (126 MHz, $CDCl_3$) δ 156.5, 152.4, 148.5, 124.6, 123.4, 122.1, 118.6, 103.9, 86.5, 28.1, 27.6, 25.9, 18.5, -4.2. MS (ESI) *m/z* = 349.3 $[M+H]^+$.



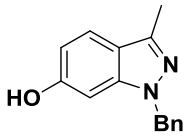
3.19

***tert*-Butyl 6-((*tert*-butyldimethylsilyl)oxy)-2*H*-indazole-2-carboxylate (3.19)**. ^{1}H NMR (500 MHz, Chloroform-*d*) δ 8.04 (s, 1H), 7.60 (s, 2H), 7.53 (dd, *J* = 8.6, 0.7 Hz, 2H), 6.83 (dd, *J* = 8.6, 2.1 Hz, 1H), 4.09 (q, *J* = 7.2 Hz, 1H), 1.69 (s, 10H), 0.99 (s, 10H), 0.23 (s, 5H). ^{13}C NMR (126 MHz, $CDCl_3$) δ 157.3, 149.5, 141.3, 139.6, 121.8, 120.8, 118.6, 105.2, 84.8, 28.4, 25.8, 18.4, -4.2. MS (ESI) *m/z* = 349.4 $[M+H]^+$.



3.20

***tert*-Butyl 6-hydroxy-1*H*-indazole-1-carboxylate (3.20).** To a stirred solution of *tert*-butyl 6-((*tert*-butyldimethylsilyl)oxy)-1*H*-indazole-1-carboxylate (**3.18**) (0.136 g, 0.39 mmol) in 10 mL anhydrous THF at 0˚C, was added TBAF solution (1 M in THF, 0.58 mL, 0.58 mmol) dropwise. The reaction was allowed to warm to room temperature overnight. The reaction mixture was poured into 20 mL saturated NH_4Cl (aq) and then extracted with ethyl acetate (2 × 20 mL). The organic layers were combined and concentrated *in vacuo*. The residue was purified by column chromatography (hexanes : ethyl acetate = 7:3) to yield 82 mg product as a white solid (90%). ^{1}H NMR (500 MHz, Chloroform-*d*) δ 8.09 (s, 1H), 7.68 (s, 1H), 7.54 (d, *J* = 8.6 Hz, 1H), 6.96 (dd, *J* = 8.6, 2.1 Hz, 1H), 1.61 (s, 6H). ^{13}C NMR (126 MHz, $CDCl_3$) δ 159.0, 149.5, 141.4, 139.8, 122.3, 119.7, 115.2, 100.3, 85.5, 28.3. MS (ESI) *m/z* = 234.9 $[M+H]^+$.

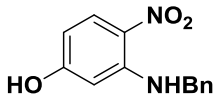


3.23b

1-Benzyl-3-methyl-1*H*-indazol-6-ol (3.23b). To a stirred solution of 1.0 g 1-(2-fluoro-4-hydroxyphenyl)ethan-1-one (6.5 mmol, 1 eq) in 20 mL xylenes was added 3.1 g $BnNHNH_2$•HCl (19.5 mmol, 3 eq) and 3.2 g KOAc (32.5 mmol, 5 eq). The reaction was heated at 140˚C for 10 d. The reaction was cooled to room temperature and poured into 50 mL 2 M HCl (aq) and extracted with ethyl acetate (2 × 30 mL). The organic layers were combined and concentrated under reduced pressure. Crude product was purified by column chromatography (hexanes : ethyl acetate = 8:2) to afford 0.8 g product as a light brown

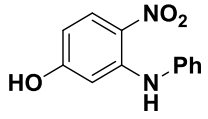
solid (52%).[1]H NMR (500 MHz, Chloroform-*d*) δ 7.48 (d, *J* = 8.7 Hz, 1H), 7.28 – 7.21 (m, 3H), 7.18 – 7.11 (m, 2H), 6.69 (dd, *J* = 8.6, 2.0 Hz, 1H), 6.59 (d, *J* = 2.0 Hz, 1H), 5.38 (s, 2H), 2.52 (s, 3H). [13]C NMR (125 MHz, $CDCl_3$) δ ppm 173.3, 156.0, 142.2, 142.0, 137.2, 128.9, 127.3, 127.3, 121.9, 111.8, 93.8, 52.7, 12.0. MS (ESI) *m/z* = 239.4 $[M+H]^+$.

General procedure for 3-alkylamino-4-nitrophenols. To a sealed vial was added 3-fluoro-4-nitrophenol (1 eq.), amine (1.5 eq) and 10 mL EtOH. The reaction was heated at 75˚C for 24 h. Upon completion, the reaction mixture was cooled to room temperature and then poured into 50 mL cold 2 M HCl (aq) while stirring. Yellow to orange precipitate was formed. The mixture was cooled in an ice bath and the product was collected by filtration.



3.25

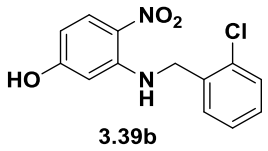
3-(Benzylamino)-4-nitrophenol (3.25). Orange solid; yield 98%; m.p. 184.2-188.0˚C. [1]H NMR (500 MHz, DMSO) δ ppm 10.75 (s, 1H), 8.77 (t, *J*=5.8 Hz, 1H), 7.99 (d, *J*=10 Hz, 1H), 7.34 (m, 4H), 7.24 (m, 1H), 6.16 (dd, *J*=2.3 Hz, *J*=9.8 Hz, 1H), 6.11 (d, *J*=2 Hz, 1H), 4.53 (d, *J*=6 Hz, 2H). [13]C NMR (125 MHz, DMSO) δ ppm 165.4, 148.3, 139.0, 129.7, 129.3, 127.8, 127.6, 125.7, 107.1, 98.4, 46.5, 40.2. MS (ESI) *m/z* = 245.0 $[M+H]^+$.



3.39a

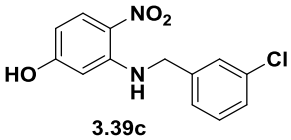
4-Nitro-3-(phenylamino)phenol (3.39a). Orange solid; yield 46%, m.p. 165.3—167.0˚C. [1]H NMR (500 MHz, DMSO-d_6) δ 10.72 (s, 1H), 9.56 (s, 1H), 8.03 (d, *J* = 9.4 Hz,

1H), 7.43 (t, *J* = 7.9 Hz, 2H), 7.36 – 7.30 (m, 2H), 7.26 – 7.19 (m, 1H), 6.44 (d, *J* = 2.4 Hz, 1H), 6.28 (dd, *J* = 9.4, 2.4 Hz, 1H). ^{13}C NMR (126 MHz, DMSO) δ 200.9, 165.0, 145.7, 139.2, 130.0, 129.5, 126.6, 125.8, 125.1, 108.5, 99.5. MS (ESI) *m/z* = 230.9 $[M+H]^+$.



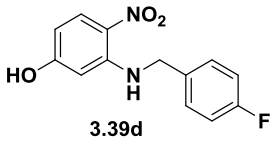
3.39b

3-((2-Chlorobenzyl)amino)-4-nitrophenol (3.39b). Orange solid; yield 78%, m.p. 202.2—203.1˚C. ^{1}H NMR (500 MHz, DMSO-d_6) δ 10.70 (s, 1H), 8.73 (t, *J* = 6.1 Hz, 1H), 8.00 (dd, *J* = 9.5, 1.8 Hz, 1H), 7.48 (dq, *J* = 7.9, 2.7 Hz, 1H), 7.37 (s, 0H), 7.30 (s, 2H), 7.30 (s, 1H), 6.16 (dd, *J* = 9.3, 2.3 Hz, 1H), 5.98 (d, *J* = 2.4 Hz, 1H), 4.59 (d, *J* = 6.1 Hz, 2H), 2.06 (d, *J* = 2.0 Hz, 2H). ^{13}C NMR (126 MHz, DMSO) δ 165.3, 147.7, 135.6, 132.5, 129.9, 129.6, 129.3, 128.8, 127.9, 125.7, 107.0, 97.9, 44.3, 31.1. MS (ESI) *m/z* = 279.0 $[M+H]^+$.



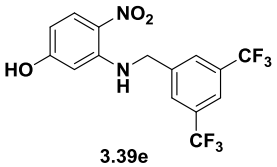
3.39c

3-((3-Chlorobenzyl)amino)-4-nitrophenol (3.39c). Orange solid, yield 82%, m.p. 208.1-209.2˚C. ^{1}H NMR (500 MHz, DMSO-d_6) δ 10.71 (s, 1H), 8.77 (t, *J* = 6.1 Hz, 1H), 7.98 (d, *J* = 9.4 Hz, 1H), 7.42 – 7.33 (m, 2H), 7.33 – 7.27 (m, 2H), 6.15 (dd, *J* = 9.4, 2.4 Hz, 1H), 6.05 (d, *J* = 2.4 Hz, 1H), 4.54 (d, *J* = 6.1 Hz, 2H). ^{13}C NMR (126 MHz, DMSO) δ 165.2, 147.8, 141.7, 133.7, 130.9, 129.5, 127.5, 127.2, 125.9, 125.6, 106.9, 98.1, 45.7. MS (ESI) *m/z* = 279.3 $[M+H]^+$.



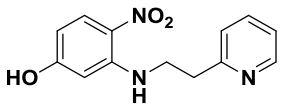
3.39d

3-((4-Fluorobenzyl)amino)-4-nitrophenol (3.39d). Orange solid, yield 67%; m.p. 198.7—199.7˚C. ^{1}H NMR (500 MHz, DMSO-d_6) δ 10.70 (s, 1H), 8.73 (t, J = 6.0 Hz, 1H), 7.97 (d, J = 9.4 Hz, 1H), 7.41 – 7.34 (m, 2H), 7.20 – 7.12 (m, 2H), 6.15 (dd, J = 9.4, 2.4 Hz, 1H), 6.08 (d, J = 2.4 Hz, 1H), 4.50 (d, J = 5.9 Hz, 2H). ^{13}C NMR (126 MHz, DMSO) δ 165.2, 162.7, 160.8, 147.9, 135.0, 134.9, 129.5, 129.4, 129.3, 125.6, 115.8, 115.7, 106.8, 98.1, 45.5, 31.1. MS (ESI) *m/z* = 263.9 $[M+H]^+$.



3.39e

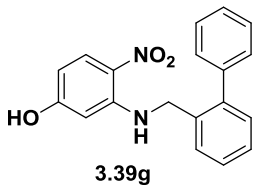
3-((3,5-Bis(trifluoromethyl)benzyl)amino)-4-nitrophenol (3.39e). Yellow solid; yield 74%; m.p. 184.0-184.8˚C. ^{1}H NMR (500 MHz, DMSO-d_6) δ 10.71 (s, 1H), 8.86 (t, J = 6.4 Hz, 1H), 8.05 (s, 2H), 8.02 – 7.96 (m, 2H), 6.17 (dd, J = 9.2, 2.4 Hz, 1H), 6.06 (d, J = 2.4 Hz, 1H), 4.72 (d, J = 6.4 Hz, 2H). ^{13}C NMR (126 MHz, DMSO) δ 165.2, 147.5, 143.1, 130.9, 130.7, 129.6, 128.2, 125.8, 124.9, 122.7, 121.3, 106.9, 98.1, 45.4. MS (ESI) *m/z* = 381.4 $[M+H]^+$.



3.39f

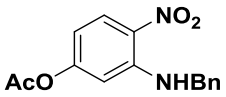
4-Nitro-3-((2-(pyridin-2-yl)ethyl)amino)phenol (3.39f). Dark purple solid; yield 42%; m.p. decomposes >220˚C. ^{1}H NMR (500 MHz, DMSO-d_6) δ 11.00 (s, 1H), 8.81 – 8.76 (m, 1H), 8.43 (td, J = 7.9, 1.6 Hz, 1H), 8.27 (s, 1H), 7.94 (dd, J = 8.8, 2.3 Hz, 2H), 7.85 (t, J = 6.8 Hz, 1H), 6.39 (d, J = 2.4 Hz, 1H), 6.23 (dd, J = 9.4, 2.3 Hz, 1H), 3.74 (t, J

= 6.9 Hz, 2H), 3.37 (t, *J* = 6.8 Hz, 2H). ^{13}C NMR (126 MHz, DMSO) δ 165.7, 154.9, 147.6, 145.6, 142.6, 129.4, 127.7, 125.5, 125.3, 107.0, 97.7, 41.8, 32.8. MS (ESI) *m/z* = 259.9 $[M+H]^+$.



3.39g

3-(([1,1'-Biphenyl]-2-ylmethyl)amino)-4-nitrophenol (2.39g). Orange solid, yield 79%; m.p. 197.3—197.9˚C. ^{1}H NMR (500 MHz, DMSO-d_6) δ 10.68 (s, 1H), 8.44 (t, *J* = 5.7 Hz, 1H), 7.94 (d, *J* = 9.4 Hz, 1H), 7.48 – 7.41 (m, 2H), 7.43 – 7.32 (m, 5H), 7.29 – 7.21 (m, 1H), 6.13 (dd, *J* = 9.4, 2.3 Hz, 1H), 5.88 (d, *J* = 2.4 Hz, 1H), 4.41 (d, *J* = 5.7 Hz, 2H). ^{13}C NMR (126 MHz, DMSO) δ 165.2, 147.7, 141.5, 140.6, 135.4, 130.5, 129.4, 129.2, 128.9, 128.2, 127.8, 127.8, 125.5, 106.8, 97.8, 44.9. MS (ESI) *m/z* = 321.5 $[M+H]^+$.

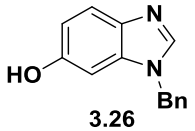


3.28

3-(Benzylamino)-4-nitrophenyl acetate (3.28). To stirred solution of 3-(benzylamino)-4-nitrophenol (**3.25**) (0.2 g, 0.892 mmol) in 10 mL pyridine at room temperature was added acetic anhydride (0.1 mL, 1.07 mmol) dropwise. The reaction was completed after 3 h. The reaction mixture was poured into 100 mL 2 M HCl aqueous solution and extracted with ethyl acetate (50 mL × 2). The organic layers were combined, washed with saturated $NaHCO_3$ and brine (50 mL each) and dried over Na_2SO_4. The solvent was removed *in vacuo* to afford the product as yellow solid (0.224g, 88%). m.p. 109.3—110.5˚C. ^{1}H NMR (500 MHz, DMSO-d_6) δ 8.70 (t, *J* = 6.1 Hz, 1H), 8.13 (d, *J* = 9.3 Hz, 1H), 7.37 – 7.30 (m, 4H), 7.30 – 7.21 (m, 1H), 6.68 (d, *J* = 2.3 Hz, 1H), 6.48 (dd,

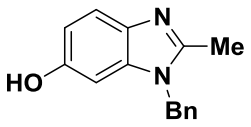
J = 9.3, 2.4 Hz, 1H), 4.58 (d, *J* = 6.1 Hz, 2H), 2.22 (s, 2H). ^{13}C NMR (126 MHz, DMSO) δ 169.0, 157.3, 146.9, 138.8, 129.7, 129.3, 128.8, 127.8, 127.6, 110.8, 107.5, 46.4, 21.6. MS (ESI) *m/z* = 287.7 $[M+H]^+$.

General procedure for the synthesis of benzo[*d*]imidazols. To a sealed tube was added 3-(R-amino)-4-nitrophenol (1 eq), NH_4Cl (10 eq) and iron powder (10 eq). An equal volume of HCOOH (88%) and *i*-PrOH was added as a solvent. The reaction was heated at 82˚C for 16 h and turned milky white. The excess HCOOH was neutralized with aqueous 2 M NaOH (aq) to pH 7. After extracting with ethyl acetate (50 mL × 2), the aqueous layer was discarded. The organic layers were combined and extracted with 2 M HCl (aq) (3 × 30 mL). The aqueous layers were combined and neutralized to pH 7 by adding 2 M NaOH (aq). A precipitate was formed, and the aqueous solution was extracted with ethyl acetate (3 × 30 mL). Organic layers were combined, washed with brine (50 mL), dried over Na_2SO_4, and evaporated under reduced pressure to afford the product as off-white to light brown solids.



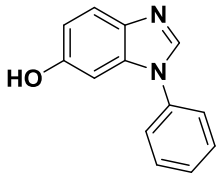
3.26

1-Benzyl-1*H*-benzo[*d*]imidazol-6-ol (3.26). Light brown solid; yield 70%. ^{1}H NMR (500 MHz, DMSO) δ ppm 8.16 (s, 1H), 7.47 (d, *J*=9 Hz, 1H), 7.28-7.34 (m, 3H), 7.24 (m, 2H), 6.79 (dd, *J*=2 Hz, *J*=9 Hz, 1H), 6.74 (d, *J*=2.5 Hz), 5.38 (s, 2H), 4.88 (s, 2H). ^{13}C NMR (125 MHz, DMSO) δ ppm 154.6, 142.4, 136.3, 136.0, 134.5, 128.8, 127.9, 127.2, 119.0, 112.6, 95.9, 47.8. MS (ESI) *m/z* = 225.1 $[M+H]^+$.



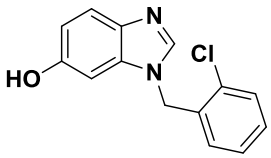
3.27

1-Benzyl-2-methyl-1*H*-benzo[*d*]imidazol-6-ol (3.27). Off-white solid; yield 35%; m.p.253.2—254.3˚C. ^{1}H NMR (500 MHz, DMSO) δ ppm 7.40 (d, *J*=9 Hz, 1H), 7.25-7.33 (m, 3H), 7.11 (m, 2H), 6.78 (dd, *J*=8.5 Hz, *J*=2.5 Hz, 1H), 6.75 (d, *J*=2 Hz, 1H), 5.362 (s, 2H), 2.53 (s, 3H). ^{13}C NMR (125 MHz, DMSO) δ ppm 154.4, 150.9, 136.1, 135.8, 133.5, 128.9, 127.8, 126.4, 117.6, 112.3, 95.9, 46.8, 12.0. MS (ESI) *m/z* = 239.3 $[M+H]^+$.



3.40a

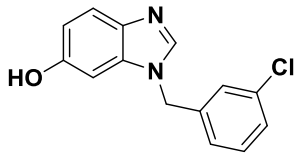
1-Phenyl-1*H*-benzo[*d*]imidazol-6-ol (3.40a). Light brown solid; yield 84%; m.p. 191.8—198.6˚C. ^{1}H NMR (500 MHz, DMSO-d_6) δ 9.38 (s, 1H), 8.30 (s, 1H), 7.64 – 7.56 (m, 4H), 7.53 (d, *J* = 8.7 Hz, 1H), 7.46 (tt, *J* = 5.5, 3.1 Hz, 1H), 6.91 (d, *J* = 2.2 Hz, 1H), 6.77 (dd, *J* = 8.7, 2.3 Hz, 1H). ^{13}C NMR (126 MHz, DMSO) δ 155.0, 142.1, 137.6, 136.7, 134.4, 130.5, 128.0, 123.9, 120.8, 112.6, 96.0. MS (ESI) *m/z* = 211.2 $[M+H]^+$.



3.40b

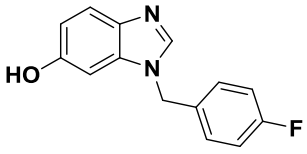
1-(2-Chlorobenzyl)-1*H*-benzo[*d*]imidazol-6-ol (3.40b). Light brown solid; yield 72%; m.p. decomposes >218˚C. ^{1}H NMR (500 MHz, DMSO-d_6) δ 9.24 (s, 1H), 8.11 (s, 1H), 7.50 (d, *J* = 7.9 Hz, 1H), 7.44 (d, *J* = 8.4 Hz, 1H), 7.32 (td, *J* = 7.6, 1.8 Hz, 1H), 7.27 (t, *J* = 7.3 Hz, 1H), 6.96 (dd, *J* = 7.6, 1.8 Hz, 1H), 6.69 (s, 2H), 6.72 – 6.65 (m, 1H), 5.46

(s, 2H). ^{13}C NMR (126 MHz, DMSO) δ 154.3, 143.6, 137.4, 135.0, 134.4, 132.6, 130.1, 130.0, 129.5, 128.0, 120.3, 112.0, 96.1, 45.9. MS (ESI) *m/z* = 258.5 $[M+H]^+$.



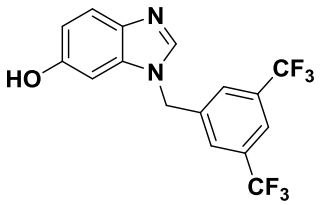
3.40c

1-(3-Chlorobenzyl)-1*H*-benzo[*d*]imidazol-6-ol (2.40c). Light brown sold; yield 79%; m.p. decomposes >252˚C. ^{1}H NMR (500 MHz, DMSO-d_6) δ 9.30 (s, 1H), 8.17 (s, 1H), 7.41 (d, *J* = 8.6 Hz, 1H), 7.39 – 7.30 (m, 1H), 7.17 (d, *J* = 6.6 Hz, 1H), 6.74 (d, *J* = 2.3 Hz, 1H), 6.67 (dd, *J* = 8.6, 2.3 Hz, 1H), 5.39 (s, 2H). ^{13}C NMR (126 MHz, DMSO) δ 154.3, 143.3, 140.0, 137.5, 134.8, 133.7, 131.0, 128.1, 127.5, 126.4, 120.3, 112.0, 96.1, 47.2. MS (ESI) *m/z* = 258.4 $[M+H]^+$.



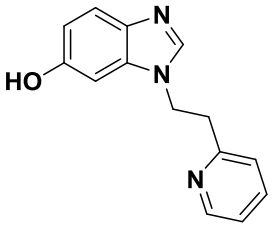
3.40d

1-(4-Fluorobenzyl)-1*H*-benzo[*d*]imidazole-6-ol (3.40d). Light brown solid; yield 72%; m.p. decomposes >183˚C. ^{1}H NMR (500 MHz, DMSO-d_6) δ 9.21 (s, 1H), 8.15 (s, 1H), 7.41 (d, *J* = 8.6 Hz, 1H), 7.30 (dd, *J* = 8.4, 5.4 Hz, 2H), 7.15 (t, *J* = 8.7 Hz, 2H), 6.74 (d, *J* = 2.2 Hz, 1H), 6.66 (dd, *J* = 8.6, 2.2 Hz, 1H), 5.35 (s, 2H). ^{13}C NMR (126 MHz, DMSO) δ 163.0, 161.0, 154.2, 143.2, 137.6, 134.8, 133.7, 133.7, 129.8, 129.8, 120.2, 116.0, 115.8, 111.9, 96.2, 47.2. MS (ESI) *m/z* = 242.4 $[M+H]^+$.



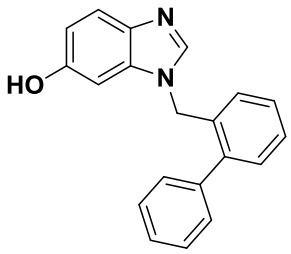
3.40e

1-(3,5-Bis(trifluoromethyl)benzyl)-1*H*-benzo[*d*]imidazol-6-ol (3.40e). Brown solid; yield 58%; m.p. >250˚C. ^{1}H NMR (500 MHz, DMSO-d_6) δ 9.29 (s, 1H), 8.26 (s, 1H), 8.04 (s, 1H), 7.97 (s, 2H), 7.43 (d, *J* = 8.6 Hz, 1H), 6.80 (s, 1H), 6.68 (d, *J* = 8.7 Hz, 1H), 5.59 (s, 2H). ^{13}C NMR (126 MHz, DMSO) δ 154.6, 143.5, 141.3, 137.7, 135.0, 131.6, 131.4, 131.1, 130.8, 128.9, 127.1, 124.9, 122.8, 122.3, 120.6, 112.4, 96.2, 47.0. MS (ESI) *m/z* = 360.8 $[M+H]^+$.



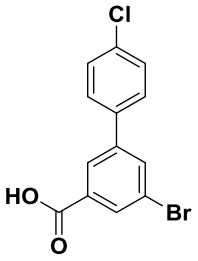
3.40f

1-(2-(Pyridin-2-yl)ethyl)-1*H*-benzo[*d*]imidazol-6-ol (3.40f). Dark purple solid; yield 40%; m.p. decomposes >156˚C. ^{1}H NMR (500 MHz, DMSO-d_6) δ 9.29 (s, 1H), 8.53 – 8.48 (m, 2H), 7.82 (s, 1H), 7.63 (td, *J* = 7.7, 1.9 Hz, 1H), 7.37 (d, *J* = 8.6 Hz, 1H), 7.20 (dd, *J* = 6.5, 4.8 Hz, 0H), 7.15 (d, *J* = 7.8 Hz, 1H), 6.85 (d, *J* = 2.2 Hz, 1H), 6.67 (dd, *J* = 8.6, 2.3 Hz, 1H), 4.50 (t, *J* = 7.1 Hz, 2H), 3.21 (t, *J* = 7.1 Hz, 2H). ^{13}C NMR (126 MHz, DMSO) δ 158.5, 154.4, 149.8, 143.1, 137.5, 137.3, 135.2, 124.1, 122.5, 120.3, 112.0, 96.0, 44.1, 37.9. MS (ESI) *m/z* = 240.3 $[M+H]^+$.



3.40g

1-([1,1'-Biphenyl]-2-ylmethyl)-1*H*-benzo[*d*]imidazol-6-ol (3.40g). Light brown solid; yield 66%; m.p. 199.5—203.3˚C. ^{1}H NMR (500 MHz, DMSO-d_6) δ 9.16 (s, 1H), 7.69 (s, 1H), 7.51 – 7.44 (m, 1H), 7.45 – 7.31 (m, 5H), 7.32 – 7.24 (m, 1H), 6.91 (d, *J* = 7.6 Hz, 1H), 6.63 (dt, *J* = 8.7, 1.8 Hz, 1H), 6.42 (t, *J* = 1.8 Hz, 1H), 5.31 (d, *J* = 2.3 Hz, 2H). ^{13}C NMR (126 MHz, DMSO) δ 154.1, 143.2, 141.3, 140.3, 137.4, 134.9, 134.1, 130.6, 129.4, 129.0, 128.2, 128.2, 128.0, 127.7, 120.2, 111.9, 96.0, 46.3. MS (ESI) *m/z* = 301.3 $[M+H]^+$.

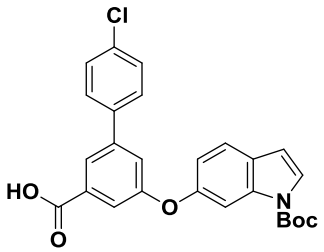


3.7

5-Bromo- 4'-chloro-[1,1'-biphenyl]-3-carboxylic acid (3.7). To a sealed tube with a magnetic stirrer was added 3-iodo-5-bromo-benzoic acid (1.514 g mg, 4.647 mmol), (4-chlorophenyl)boronic acid (0.707 g, 4.647 mmol), Cs_2CO_3 (1.651 mg, 5.112 mmol) and $Pd(PPh_3)_4$ (0.268 g, 0.232 mmol.). A mixed solvent (toluene : ethanol : water = 9:3:1, 50 mL) was added. The reaction was heated at 80℃ for 24 h. Upon completion, the reaction mixture was poured into 100 mL 2 M HCl and extracted with ethyl acetate (50 mL × 2). The organic layers were combined and filtered through a cotton plug, concentrated under

reduced pressure. The residue was purified by column chromatography (hexanes : ethyl acetate = 8:2) to afford the product as off-white to light orange solid (0.93 g, 64%). ^{1}H NMR (500 MHz, DMSO) δ ppm 8.12 (s, 2H), 8.01 (s, 1H), 7.76 (d, *J* = 8.5 Hz, 2H), 7.52 (d, *J* = 8.5 Hz, 2H). ^{13}C NMR (125 MHz, DMSO) δ ppm 166.5, 142.2, 137.2, 134.2, 134.1, 131.5, 129.8, 138.6, 127.0, 123.2. MS (ESI) *m/z* = 311.2 $[M + H]^{+}$.

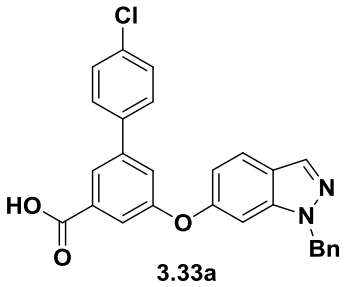
General procedure for Ullman reaction. To a sealed vial under argon was added aryl bromide (1.0 eq.), phenol (1.2 eq.), Cs_2CO_3 (3 eq.), CuI (0.1 eq.), N,N-dimethyl glycine hydrochloride (0.3 eq.) and 10 mL anhydrous dioxane. The reaction was heated at 95℃ for 16 h. Upon completion, the solvent was evaporated under reduced pressure. The residue was broken up in 20 mL 2 M HCl (aq) and extracted with ethyl acetate (15 mL × 2). The organic layers were combined and washed with saturated Na_2CO_3 (aq) and brine (20 mL each). The organic layer was then evaporated under reduced pressure. Crude product was purified by column chromatography (hexanes : ethyl acetate = 8:2) or crystalized from adding hexanes to ethyl acetate solution to afford desired product.



3.9

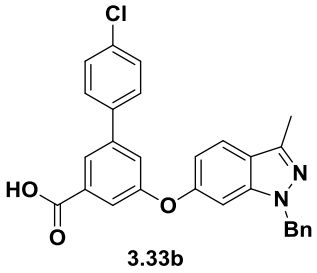
5-((1-(*tert*-Butoxycarbonyl),-1*H*-indol-6-yl),oxy),-4'-chloro-[1,1'- biphenyl]-3-carboxylic acid (3.9). as off-white solid (0.15 g, 65%). ^{1}H NMR (500 MHz, DMSO) δ ppm 7.93 (br, 1H), 7.62-7.71 (m, 6H), 7.49 (d, 2H, *J* = 7.5 Hz), 7.43 (s, 1H), 7.08 (d, 1H, *J* = 8.5 Hz), 6.70 (d, 1H, *J* = 3.5 Hz), 1.47 (s, 9H), ^{13}C NMR (125 MHz, DMSO) δ ppm 167.2, 158.8, 154.3, 149.6, 142.0, 138.0, 135.7, 134.1, 133.9, 129.7, 129.4, 128.3, 127.4,

127.0, 123.1, 122.9, 121.8, 118.2, 116.0, 108.1, 106.1, 84.7, 28.1. MS (ESI) *m/z* = 464.2 $[M+H]^+$.



3.33a

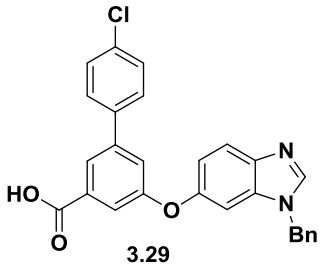
5-((1-Benzyl-1*H*-indazol-6-yl)oxy)-4'-chloro-[1,1'-biphenyl]-3-carboxylic acid (3.33a). Off-white solid; yield 58%. ^{1}H NMR (500 MHz, DMSO-d_6) δ 13.22 (s, 1H), 8.10 (d, *J* = 1.0 Hz, 1H), 7.93 (t, *J* = 1.6 Hz, 1H), 7.81 (d, *J* = 8.7 Hz, 1H), 7.72 – 7.66 (m, 2H), 7.58 (t, *J* = 2.1 Hz, 1H), 7.55 – 7.44 (m, 3H), 7.41 (dd, *J* = 2.4, 1.3 Hz, 1H), 7.27 – 7.15 (m, 5H), 6.95 (dd, *J* = 8.7, 2.1 Hz, 1H), 5.58 (s, 2H). ^{13}C NMR (126 MHz, DMSO) δ 166.9, 158.5, 155.2, 141.7, 140.5, 137.8, 133.9, 133.6, 129.5, 129.1, 128.9, 127.9, 123.2, 122.6, 121.3, 121.1, 117.6, 114.8, 100.2, 52.2. MS (ESI) *m/z* = 455.0 $[M+H]^+$.



3.33b

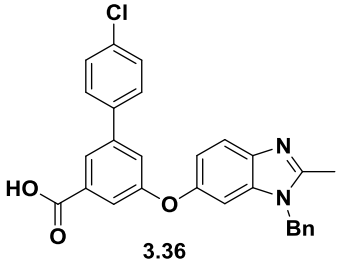
5-((1-Benzyl-3-methyl-1*H*-indazol-6-yl)oxy)-4'-chloro-[1,1'-biphenyl]-3-carboxylic acid (3.33b). Off-white solid; yield 60%. ^{1}H NMR (500 MHz, Chloroform-*d*) δ 8.03 (t, *J* = 1.6 Hz, 1H), 7.66 (dd, *J* = 2.4, 1.4 Hz, 1H), 7.64 (d, *J* = 8.7 Hz, 1H), 7.49 – 7.46 (m, 2H), 7.42 – 7.38 (m, 3H), 7.25 – 7.12 (m, 5H), 6.89 (dd, *J* = 8.7, 2.0 Hz, 1H), 6.86 (d, *J* = 1.9 Hz, 1H), 5.45 (s, 2H), 2.59 (s, 3H). ^{13}C NMR (125 MHz, $CDCl_3$) δ ppm 170.0, 158.3, 156.1, 142.4, 142.4, 141.4, 137.9, 136.9, 134.6, 132.0, 129.4, 128.9, 128.6, 127.9,

127.4, 123.6, 122.4, 122.2, 120.9, 119.0, 113.8, 98.9, 52.9, 12.1. MS (ESI) *m/z* = 468.9 $[M+H]^+$.



3.29

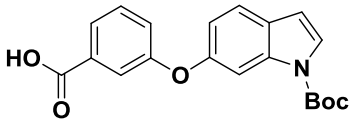
5-((1-Benzyl-1*H*-benzo[*d*]imidazol-6-yl)oxy)-4'-chloro-[1,1'-biphenyl]-3-carboxylic acid (3.29). Off-white solid; yield 67%. ^{1}H NMR (500 MHz, DMSO-d_6) δ 7.93 (d, *J* = 1.7 Hz, 1H), 7.73 – 7.66 (m, 2H), 7.55 (d, *J* = 2.1 Hz, 1H), 7.52 (d, *J* = 8.2 Hz, 3H), 7.36 (q, *J* = 2.4, 1.9 Hz, 4H), 7.26 (dt, *J* = 20.0, 7.4 Hz, 5H), 5.60 (s, 2H). ^{13}C NMR (126 MHz, DMSO) δ 167.1, 158.9, 153.6, 141.9, 138.0, 136.0, 134.1, 134.0, 133.9, 129.7, 129.4, 129.4, 129.4, 128.9, 128.5, 122.8, 121.1, 117.6, 49.5. MS (ESI) *m/z* = 455.5 $[M+H]^+$.



3.36

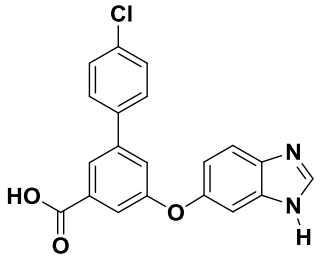
5-((1-Benzyl-2-methyl-1*H*-benzo[*d*]imidazol-6-yl)oxy)-4'-chloro-[1,1'-biphenyl]-3-carboxylic acid (3.36). Off-white solid; yield 52%. ^{1}H NMR (500 MHz, DMSO-d_6) δ 7.94 (dt, *J* = 2.9, 1.4 Hz, 1H), 7.84 (d, *J* = 9.0 Hz, 1H), 7.69 (dd, *J* = 8.5, 2.5 Hz, 2H), 7.65 (d, *J* = 3.1 Hz, 1H), 7.57 (qd, *J* = 2.7, 1.8, 1.2 Hz, 1H), 7.52 (dd, *J* = 8.6, 2.5 Hz, 2H), 7.36 (dp, *J* = 2.4, 1.2 Hz, 1H), 7.35 – 7.22 (m, 6H), 5.64 (d, *J* = 2.4 Hz, 2H), 2.81 (d, *J* = 2.4 Hz, 3H). ^{13}C NMR (126 MHz, DMSO) δ 167.1, 158.7, 154.1, 142.0, 137.9, 135.1, 134.2, 133.9, 129.7, 129.5, 129.4, 128.9, 128.0, 123.0, 121.3, 118.9, 117.7, 116.9,

104.4, 48.2, 12.8. MS (ESI) *m/z* = 469.4 $[M+H]^+$.



3.48

3-((1-(*tert*-Butoxycarbonyl)-1*H*-indol-6-yl)oxy)benzoic acid (3.48). White solid; yield 63%. ^{1}H NMR (500 MHz, Chloroform-*d*) δ 11.47 (s, 1H), 7.82 (dt, *J* = 7.8, 1.2 Hz, 1H), 7.74 – 7.70 (m, 1H), 7.59 (d, *J* = 3.7 Hz, 1H), 7.53 (d, *J* = 8.4 Hz, 1H), 7.42 (t, *J* = 7.9 Hz, 1H), 7.32 – 7.27 (m, 1H), 6.99 (dd, *J* = 8.4, 2.2 Hz, 1H), 6.56 (dd, *J* = 3.7, 0.9 Hz, 1H), 1.60 (s, 9H). ^{13}C NMR (126 MHz, $CDCl_3$) δ 171.9, 158.7, 154.0, 149.8, 131.2, 130.0, 127.3, 126.3, 124.7, 124.0, 122.0, 119.8, 115.7, 107.3, 107.0, 84.2, 28.3. MS (ESI) *m/z* = 354.1 $[M+H]^+$.



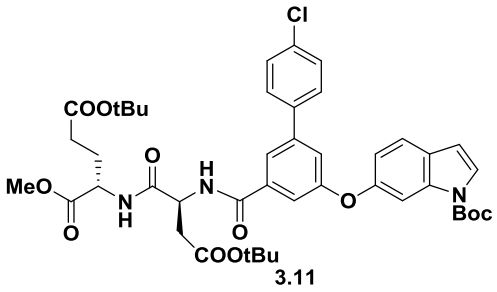
3.32

5-((1*H*-Benzo[*d*]imidazol-6-yl)oxy)-4'-chloro-[1,1'-biphenyl]-3-carboxylic acid (3.32). To a stirred solution of 150 mg 5-((1-benzyl-1*H*-benzo[*d*]imidazol-6-yl)oxy)-4'-chloro-[1,1'-biphenyl]-3-carboxylic acid (0.33 mmol, 1eq) in 2 mL DMSO, was added 296 mg potassium *tert*-butyl oxide (2.638 mmol, 8 eq) in small portions. Oxygen (or air) is bubbled to the solution through a needle. Upon completion (monitored by TLC, approx. 20 min), the reaction mixture was poured to 20 mL ice cooled 1 M HCl aqueous solution. The resulting milky solution was further stirred in ice bath for 10 min allowing full precipitation. The off-white solid was filtered and dried in oven (120˚C) for 1 h to afford

the title compound as light brown solid; yield 85%. ^{1}H NMR (500 MHz, DMSO-d_6) δ 13.16 (s, 1H), 8.40 (s, 1H), 7.87 (s, 1H), 7.70 (d, J = 8.0 Hz, 2H), 7.56 (s, 1H), 7.50 (d, J = 7.8 Hz, 2H), 7.34 (d, J = 12.5 Hz, 2H), 7.05 (d, J = 7.3 Hz, 1H). ^{13}C NMR (126 MHz, DMSO) δ 166.9, 159.6, 151.5, 141.5, 137.9, 133.7, 133.5, 129.5, 129.1, 122.0, 120.5, 116.7, 116.0. MS (ESI) m/z = 365.3 $[M+H]^+$.

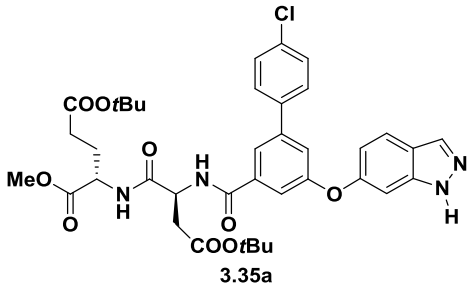
Oxidative debenzylation for the synthesis of 3.34a-b and 3.37. The procedure was adopted from the synthesis of **3.32**. Starting materials decomposed rapidly under room temperature. Therefore a mixture of solvents was used (THF : DMSO = 1:1) to allow the reactions to run at 10˚C. The reactions showed no further conversion after 30 min. A mixture of product and starting material was isolated and was used without further purification.

General procedure for peptide coupling: At 0 ˚C, to a suspension of carboxylic acid (1 equiv), amine (HCl salt, 1 equiv), EDC•HCl (1.2 equiv), and HOAt (1 equiv) in anhydrous dichloromethane was added triethylamine (2.5 equiv) dropwise. A clear yellow solution was formed. The reaction was kept at 0 ˚C for 4–6 h until TLC showed no starting material left. The reaction was concentrated under reduced pressure. Residue was broken up in ethyl acetate and ice-cold 2 M HCl (aq). The organic layer was washed with equal volume of saturated $NaHCO_3$ (aq) and brine. The organic layer was dried over Na_2SO_4 and then evaporated under reduced pressure. Column chromatography was used to purify the final product.



3.11

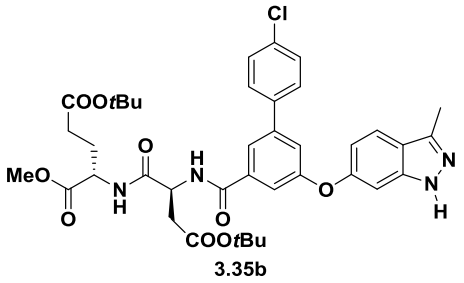
(*S*),-5-*tert*-Butyl 1-methyl 2-((*S*),-4-(*tert*-butoxy),-2-(5- ((1-(*tert*-butoxycarbonyl),-1*H*-indol-6-yl),oxy),-4'- chloro-[1,1'-biphenyl]-3-ylcarboxamido),-4- oxobutanamido),pentanedioate (3.11). Colorless syrup, yield, 90%. ^{1}H NMR (500 MHz, DMSO) δ ppm 7.87 (br, 1H), 7.69 (d, 1H, *J* = 1.5 Hz), 7.63 (d, 1H, *J* = 7.5 Hz), 7.56 (s, 1H), 7.52 (dd, 1H, *J* = 9.0 Hz, *J* = 3 Hz), 7.48 (m, 2H), 7.45 (s, 1H), 7.43 (s, 1H), 7.37 (m, 2H), 7.32 (s, 1H), 6.99 (td, 1H, *J* = 8.0 Hz, *J* = 2.5 Hz), 6.54 (d, 1H, *J* = 3.0 Hz), 4.98 (m, 1H), 4.55 (m, 1H), 3.71 (s, 3H), 3.00 (dd, 1H, *J* = 17.0 Hz, *J* = 4.0 Hz), 2.65 (dd, 1H, *J* = 17.3 Hz *J* = 7.3 Hz), 2.28 (m, 2H), 2.13 (m, 1H), 1.96 (m, 1H), 1.56 (s, 9H), 1.43 (s, 9H), 1.34 (s, 9H). 13C NMR (125 MHz, DMSO) δ ppm 172.2, 172.0, 171.9, 170.8, 166.6, 159.4, 153.8, 148.7, 142.4, 138.3, 136.2, 134.4, 129.3, 128.7, 127.4, 126.4, 122.1, 120.2, 120.0, 116.1, 115.7, 107.3, 107.2, 84.2, 82.3, 81.0, 52.7, 52.3, 49.8, 37.3, 31.5, 28.3, 28.2, 27.2. MS (ESI) *m/z* = 834.6 $[M+H]^{+}$.



3.35a

5-(*tert*-Butyl) 1-methyl ((*S*)-2-(5-((1*H*-indazol-6-yl)oxy)-4'-chloro-[1,1'-biphenyl]-3-carboxamido)-4-(*tert*-butoxy)-4-oxobutanoyl)-L-glutamate (3.35a).

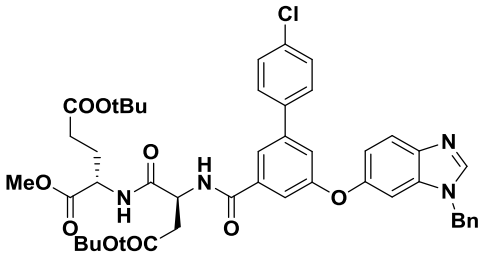
Colorless syrup; yield 40%. ^{1}H NMR (500 MHz, Chloroform-*d*) δ 8.03 (s, 1H), 7.83 (d, *J* = 7.7 Hz, 1H), 7.77 (t, *J* = 1.6 Hz, 1H), 7.72 (d, *J* = 8.7 Hz, 1H), 7.52 – 7.47 (m, 2H), 7.45 (d, *J* = 7.8 Hz, 1H), 7.41 – 7.37 (m, 2H), 7.04 (d, *J* = 2.0 Hz, 1H), 6.97 (dd, *J* = 8.7, 2.0 Hz, 1H), 5.00 (td, *J* = 7.4, 4.1 Hz, 1H), 4.52 (td, *J* = 7.9, 5.2 Hz, 1H), 3.72 (s, 3H), 2.98 (dd, *J* = 17.1, 4.1 Hz, 1H), 2.66 (dd, *J* = 17.1, 7.1 Hz, 1H), 2.34 – 2.20 (m, 2H), 2.18 – 2.09 (m, 1H), 1.99 – 1.87 (m, 1H), 1.43 (s, 9H), 1.38 (s, 9H). ^{13}C NMR (126 MHz, $CDCl_3$) δ 172.2, 172.0, 171.9, 170.9, 166.5, 158.1, 156.7, 142.7, 138.0, 136.3, 134.6, 129.3, 128.6, 121.1, 121.0, 116.9, 115.1, 95.0, 82.4, 81.2, 60.6, 52.7, 52.2, 49.9, 37.3, 31.5, 28.2, 28.2, 27.3, 14.4. MS (ESI) *m/z* = 735.2 $[M+H]^+$.



3.35b

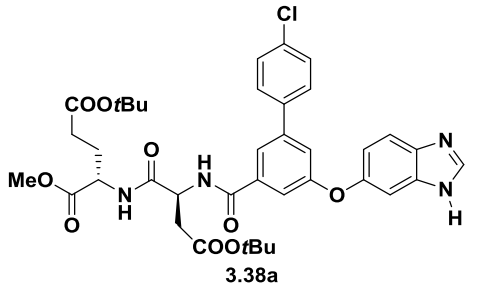
5-(*tert*-Butyl) 1-methyl ((*S*)-4-(*tert*-butoxy)-2-(4'-chloro-5-((3-methyl-1*H*-indazol-6-yl)oxy)-[1,1'-biphenyl]-3-carboxamido)-4-oxobutanoyl)-L-glutamate (3.35b). Colorless syrup; yield 34%. ^{1}H NMR (500 MHz, $CDCl_3$) δ ppm 7.82 (d, *J*=7.7 Hz, 1H) 7.76 (t, *J*=1.4 Hz, 1H) 7.62 (d, *J*=8.7 Hz, 1H) 7.49 (m, 3H) 7.44 (d, *J*=7.7 Hz, 1H) 7.38 (d, *J*=8.5 Hz, 2H) 7.36 (m, 1H) 6.98 (d, *J*=1.9 Hz, 1H) 6.92 (dd, *J*=1.9 Hz, *J*=8.7 Hz, 1H) 5.00 (dq, *J*=4.4 Hz, *J*=7.3 Hz, 1H) 4.54 (m, 1H) 3.71 (s, 3H) 2.98 (dd, *J*=4.1 Hz, *J*=17.1 Hz, 1H) 2.66 (dd, *J*=7.2 Hz, *J*=17.1 Hz, 1H) 2.56 (s, 3H) 2.26 (m, 2H) 2.13 (m, 1H) 1.94 (m, 1H) 1.43 (s, 9H) 1.38 (s, 9H). ^{13}C NMR (125 MHz, $CDCl_3$) δ ppm 172.2, 172.0, 171.9, 170.9, 166.5, 158.2, 156.8, 142.7, 138.0, 136.3, 134.6, 129.3, 128.6, 122.1, 121.0, 120.7, 116.8, 114.3, 99.0, 82.4, 81.1, 52.7, 52.2, 49.9, 37.3, 31.5, 28.2, 27.3, 14.4, 12.0. MS (ESI)

m/z = 749.4 $[M+H]^+$.



3.30

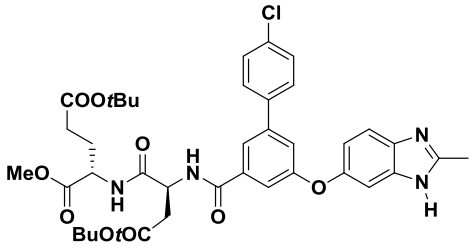
5-(*tert*-Butyl) 1-methyl ((*S*)-2-(5-((1-benzyl-1*H*-benzo[*d*]imidazol-6-yl)oxy)-4'-chloro-[1,1'-biphenyl]-3-carboxamido)-4-(*tert*-butoxy)-4-oxobutanoyl)-L-glutamate (3.30). Colorless syrup; yield 62%. ^{1}H NMR (500 MHz, Chloroform-*d*) δ 8.14 (s, 1H), 7.70 (dt, *J* = 3.2, 1.6 Hz, 1H), 7.66 (dd, *J* = 7.7, 4.6 Hz, 1H), 7.49 – 7.39 (m, 3H), 7.39 – 7.34 (m, 2H), 7.31 – 7.24 (m, 3H), 7.21 (ddd, *J* = 4.0, 2.3, 1.6 Hz, 1H), 7.19 – 7.15 (m, 2H), 7.05 (dd, *J* = 8.7, 2.2 Hz, 1H), 7.00 (s, 1H), 5.32 (s, 2H), 4.98 (td, *J* = 7.3, 4.0 Hz, 1H), 4.55 (m, 1H), 3.72 (d, *J* = 1.3 Hz, 3H), 3.01 (dd, *J* = 17.1, 3.9 Hz, 1H), 2.66 (dd, *J* = 17.2, 7.3 Hz, 1H), 2.36 – 2.21 (m, 2H), 2.18 – 2.09 (m, 1H), 2.00 – 1.89 (m, 1H), 1.44 (s, 9H), 1.37 (s, 9H). ^{13}C NMR (126 MHz, $CDCl_3$) δ 172.1, 172.0, 171.9, 170.8, 166.5, 159.1, 153.0, 142.5, 138.1, 138.1, 136.4, 134.9, 134.5, 129.4, 129.3, 128.7, 128.6, 128.6, 127.5, 127.5, 121.3, 120.4, 119.7, 116.1, 116.0, 101.6, 82.4, 82.3, 81.1, 52.7, 52.7, 52.3, 52.2, 49.8, 49.5, 37.4, 31.5, 31.5, 28.3, 28.2, 27.3. MS (ESI) *m/z* = 825.1 $[M+H]^+$.



3.38a

5-(*tert*-Butyl) 1-methyl ((*S*)-2-(5-((1*H*-benzo[*d*]imidazol-6-yl)oxy)-4'-chloro-

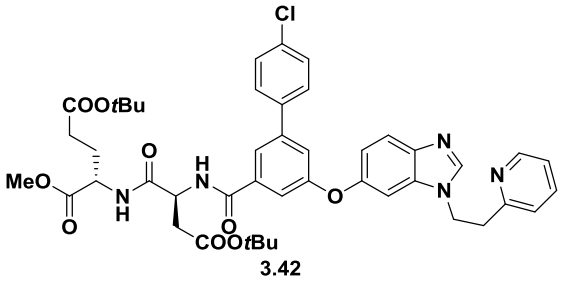
[1,1'-biphenyl]-3-carboxamido)-4-(*tert*-butoxy)-4-oxobutanoyl)-L-glutamate (3.38a). Colorless syrup; yield 27%. ^{1}H NMR (500 MHz, Chloroform-*d*) δ 8.05 (s, 1H), 7.84 (d, *J* = 7.8 Hz, 1H), 7.66 (q, *J* = 1.7 Hz, 1H), 7.60 (d, *J* = 8.5 Hz, 1H), 7.51 (d, *J* = 8.0 Hz, 1H), 7.45 – 7.38 (m, 3H), 7.36 – 7.33 (m, 2H), 7.29 – 7.24 (m, 1H), 7.20 (s, 1H), 6.99 (dd, *J* = 8.8, 2.2 Hz, 1H), 4.99 (td, *J* = 7.3, 4.3 Hz, 1H), 4.52 (ddd, *J* = 13.0, 9.2, 5.3 Hz, 1H), 3.69 (s, 3H), 2.97 (dd, *J* = 17.2, 4.1 Hz, 1H), 2.67 (dd, *J* = 17.1, 7.1 Hz, 1H), 2.34 – 2.21 (m, 2H), 2.16 – 2.07 (m, 1H), 1.93 (ddd, *J* = 12.3, 8.1, 6.2 Hz, 1H), 1.41 (s, 9H), 1.36 (s, 9H). ^{13}C NMR (126 MHz, $CDCl_3$) δ 172.3, 172.2, 172.1, 172.0, 171.7, 170.9, 166.9, 159.5, 152.3, 142.5, 142.4, 141.8, 138.1, 136.1, 134.4, 129.3, 128.6, 120.1, 119.9, 116.1, 115.6, 82.4, 81.2, 52.7, 52.7, 52.3, 49.9, 28.2, 27.2. MS (ESI) *m/z* = 735.1 $[M+H]^+$.



3.38b

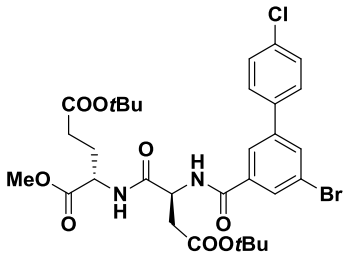
5-(*tert*-Butyl) 1-methyl ((*S*)-4-(*tert*-butoxy)-2-(4'-chloro-5-((2-methyl-1*H*-benzo[*d*]imidazol-6-yl)oxy)-[1,1'-biphenyl]-3-carboxamido)-4-oxobutanoyl)-L-glutamate (3.38b). Colorless syrup; yield 32%. ^{1}H NMR (300 MHz, $CDCl_3$) δ ppm 7.89 (d, *J*=7.7 Hz, 1H) 7.67 (t, *J*=1.4 Hz, 1H) 7.46 (m, 4H) 7.37 (m, 3H) 7.28 (m, 1H) 7.12 (d, *J*=1.9 Hz, 1H) 6.92 (dd, *J*=2.2 Hz, *J*=8.6 Hz, 1H) 4.98 (dt, *J*=4.3 Hz, *J*=7.4 Hz, 1H) 4.52 (dt, *J*=4.8 Hz, *J*=7.9 Hz, 1H) 3.71 (s, 3H) 2.97 (dd, *J*=4.1 Hz, *J*=17.3 Hz, 1H) 2.66 (dd, *J*=7.0 Hz, *J*=16.8 Hz, 1H) 2.57 (s, 3H) 2.27 (m, 2H) 2.15 (m, 1H) 1.94 (ddd, *J*=6.1 Hz, *J*=7.9 Hz, *J*=14.3 Hz, 1H) 1.42 (s, 9H) 1.37 (s, 9H). ^{13}C NMR (75 MHz, $CDCl_3$) δ ppm 172.3, 172.0, 171.8, 170.9, 166.9, 159.8, 152.1, 141.6, 142.5, 138.2, 135.9, 134.4, 129.3,

128.6, 119.9, 119.8, 115.6, 115.1, 98.8, 82.4, 81.2, 52.7, 52.3, 49.9, 37.2, 31.5, 29.9, 28.2, 27.2, 15.0. MS (ESI) *m/z* = 749.6 $[M+H]^+$.



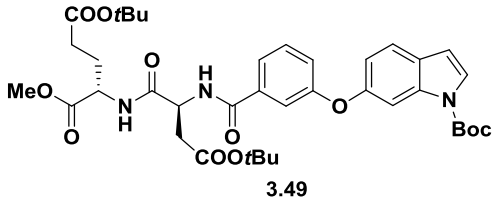
3.42

5-(*tert*-Butyl) 1-methyl ((S)-4-(*tert*-butoxy)-2-(4'-chloro-5-((1-(2-(pyridin-2-yl)ethyl)-1*H*-benzo[*d*]imidazol-6-yl)oxy)-[1,1'-biphenyl]-3-carboxamido)-4-oxobutanoyl)-L-glutamate (3.42). Colorless syrup. 53%. ^{1}H NMR (500 MHz, Chloroform-*d*) δ 8.54 (ddd, *J* = 4.9, 1.8, 0.9 Hz, 1H), 7.77 (s, 1H), 7.74 (d, *J* = 8.7 Hz, 1H), 7.70 (t, *J* = 1.6 Hz, 1H), 7.67 (d, *J* = 7.6 Hz, 1H), 7.51 – 7.42 (m, 5H), 7.41 – 7.36 (m, 2H), 7.10 (ddd, *J* = 7.6, 4.9, 1.1 Hz, 1H), 7.07 (d, *J* = 2.3 Hz, 1H), 7.01 (dd, *J* = 8.7, 2.3 Hz, 1H), 6.87 (d, *J* = 7.8 Hz, 1H), 4.98 (td, *J* = 7.4, 3.9 Hz, 1H), 4.60 (t, *J* = 6.8 Hz, 2H), 4.55 (td, *J* = 7.9, 5.1 Hz, 1H), 3.72 (s, 3H), 3.28 (t, *J* = 6.7 Hz, 2H), 3.00 (dd, *J* = 17.2, 3.9 Hz, 1H), 2.66 (dd, *J* = 17.2, 7.3 Hz, 1H), 2.37 – 2.19 (m, 2H), 2.14 (dtd, *J* = 14.9, 7.5, 5.1 Hz, 1H), 1.94 (dtd, *J* = 14.3, 8.1, 6.3 Hz, 1H), 1.44 (s, 9H), 1.37 (s, 9H). ^{13}C NMR (126 MHz, $CDCl_3$) δ 172.1, 172.0, 171.9, 170.8, 166.6, 159.6, 157.1, 152.6, 149.7, 143.9, 142.5, 138.2, 137.2, 137.1, 136.3, 134.5, 129.3, 128.6, 124.1, 124.1, 122.4, 121.5, 120.1, 119.6, 115.8, 115.6, 101.3, 82.4, 81.1, 52.7, 52.7, 52.3, 49.8, 44.7, 38.2, 37.4, 31.5, 28.3, 28.2, 27.3. MS (ESI) *m/z* = 840.5 $[M+H]^+$.



3.46

5-(*tert*-Butyl) 1-methyl ((*S*)-2-(5-bromo-4'-chloro-[1,1'-biphenyl]-3-carboxamido)-4-(*tert*-butoxy)-4-oxobutanoyl)-L-glutamate (3.46). Colorless syrup; yield 83%. ^{1}H NMR (500 MHz, Chloroform-*d*) δ 7.92 (dt, *J* = 5.2, 1.6 Hz, 2H), 7.83 (t, *J* = 1.7 Hz, 1H), 7.64 (dd, *J* = 7.6, 4.3 Hz, 1H), 7.52 (dd, *J* = 8.5, 2.2 Hz, 2H), 7.48 (d, *J* = 7.8 Hz, 1H), 7.46 – 7.40 (m, 2H), 5.00 (tt, *J* = 7.5, 3.9 Hz, 1H), 4.57 (tdd, *J* = 7.9, 5.0, 2.8 Hz, 1H), 3.74 (s, 3H), 3.04 (dd, *J* = 17.2, 3.7 Hz, 1H), 2.67 (dd, *J* = 17.2, 7.4 Hz, 1H), 2.38 – 2.21 (m, 2H), 2.17 (ddd, *J* = 14.6, 7.3, 5.1 Hz, 1H), 2.06 – 1.91 (m, 1H), 1.48 (s, 9H), 1.38 (s, 9H). ^{13}C NMR (126 MHz, $CDCl_3$) δ 172.2, 171.9, 170.7, 170.4, 165.6, 142.7, 137.4, 136.3, 134.9, 133.4, 129.4, 129.3, 129.2, 128.7, 124.8, 123.5, 82.5, 81.1, 52.7, 52.3, 49.8, 37.1, 31.5, 28.2, 27.2. MS (ESI) *m/z* = 681.5 $[M+H]^+$.

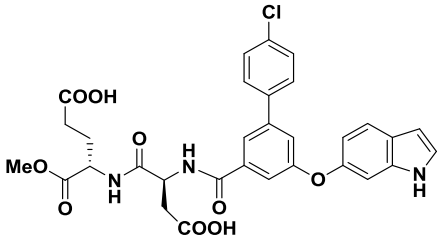


3.49

5-(*tert*-Butyl) 1-methyl ((*S*)-4-(*tert*-butoxy)-2-(3-((1-(*tert*-butoxycarbonyl)-1*H*-indol-6-yl)oxy)benzamido)-4-oxobutanoyl)-L-glutamate (3.49). Colorless syrup; yield 86%. ^{1}H NMR (500 MHz, Chloroform-*d*) δ 7.82 (s, 1H), 7.58 (dd, *J* = 12.0, 5.7 Hz, 2H), 7.53 – 7.43 (m, 3H), 7.43 – 7.34 (m, 2H), 7.18 – 7.13 (m, 1H), 6.96 (dt, *J* = 8.5, 2.0 Hz, 1H), 6.54 (d, *J* = 3.6 Hz, 1H), 4.96 (td, *J* = 7.2, 3.6 Hz, 1H), 4.60 – 4.51 (m, 1H), 3.72 (d,

J = 1.6 Hz, 3H), 3.01 (dd, *J* = 17.1, 3.8 Hz, 1H), 2.62 (dd, *J* = 17.1, 7.0 Hz, 1H), 2.36 – 2.20 (m, 2H), 2.20 – 2.09 (m, 1H), 1.99 – 1.88 (m, 1H), 1.60 (s, 9H), 1.44 (s, 10H), 1.40 (s, 9H). ^{13}C NMR (126 MHz, $CDCl_3$) δ 172.2, 172.0, 172.0, 170.8, 166.8, 153.9, 135.6, 130.1, 127.3, 126.3, 122.0, 121.3, 107.3, 84.1, 82.3, 81.0, 52.7, 52.2, 49.7, 37.2, 31.5, 28.3, 28.2, 27.3, 21.3. MS (ESI) *m/z* = 724.6 $[M+H]^+$.

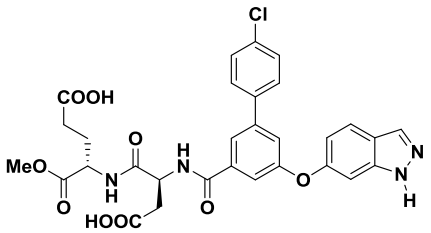
General procedure for deprotection of *tert*-Butyl ester. To a stirred solution of *tert*-butyl ester in anhydrous dichloromethane (5 mL) at 0℃ was added trifluoroacetic acid (5 mL) dropwise. The reaction was kept at 0℃ for 4-6 h until TLC showed no starting material left. Anhydrous toluene (5 mL) was added to the reaction. The solvent was removed under reduced pressure to afford the product.



3.4

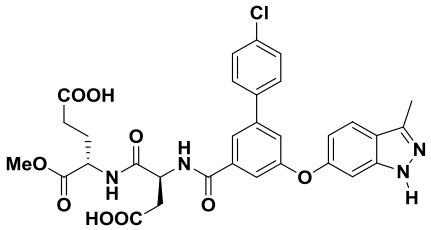
(*S*)-4-((*S*)-2-(5-((1*H*-indol-6-yl)oxy)-4'-chloro-[1,1'-biphenyl]-3-carboxamido)-3-carboxypropanamido)-5-methoxy-5-oxopentanoic acid (2.4). Off-white to light pink solid; yield, 98%; m.p. decomposes at 126.1˚C. $[\alpha]_D^{20}$ -10.4˚ (*c* 14.36, MeOH). ^{1}H NMR (500 MHz, DMSO-d_6) δ 12.18 (s, 2H), 11.01 (s, 1H), 8.62 (d, *J* = 7.7 Hz, 1H), 8.30 (d, *J* = 7.6 Hz, 1H), 7.57 (d, *J* = 7.8 Hz, 1H), 7.54 (d, *J* = 8.5 Hz, 1H), 7.48 – 7.45 (m, 1H), 7.42 (t, *J* = 7.9 Hz, 1H), 7.30 (t, *J* = 2.7 Hz, 1H), 7.11 (dd, *J* = 8.0, 2.5 Hz, 1H), 7.00 (d, *J* = 2.1 Hz, 1H), 6.75 (dd, *J* = 8.6, 2.2 Hz, 1H), 6.41 (t, *J* = 2.4 Hz, 1H), 4.77 (td, *J* = 8.5, 5.0 Hz, 1H), 4.25 (td, *J* = 8.4, 5.4 Hz, 1H), 3.59 (s, 3H), 2.72 (dd, *J* = 16.7, 4.8 Hz, 1H), 2.65 (dd, *J* = 16.7, 9.2 Hz, 1H), 2.26 (t, *J* = 8.0 Hz, 3H), 1.94 (dq, *J* = 13.4, 7.4

Hz, 1H), 1.84 – 1.73 (m, 1H). ^{13}C NMR (126 MHz, DMSO) δ 174.1, 172.5, 172.1, 171.5, 166.0, 158.7, 151.6, 136.7, 136.2, 130.1, 129.3, 128.6, 126.1, 125.7, 125.0, 122.0, 121.5, 121.0, 117.2, 112.5, 102.3, 101.6, 52.3, 51.8, 50.5, 36.4, 30.2, 26.4. HRMS (ESI) *m/z* calcd. for $C_{31}H_{28}ClN_3O_9Na^+$ $[M + Na]^+$ 644.1412, found 644.1417.



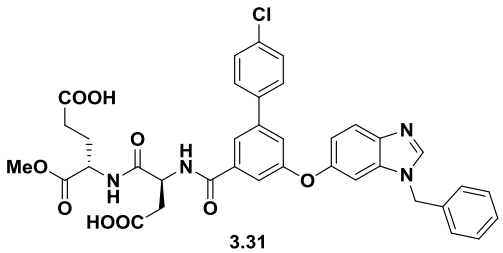
3.12

(*S*)-4-((*S*)-2-(5-((1*H*-Indazol-6-yl)oxy)-4'-chloro-[1,1'-biphenyl]-3-carboxamido)-3-carboxypropanamido)-5-methoxy-5-oxopentanoic acid (3.12). White solid; yield, 99%; m.p. decomposes at 124.5˚C. $[\alpha]_D^{20}$ -21.0˚ (*c* 3.02, MeOH). ^{1}H NMR (500 MHz, DMSO-d_6) δ 12.23 (s, 2H) 8.82 (d, *J* = 7.8 Hz, 1H), 8.37 (d, *J* = 7.5 Hz, 1H), 8.05 (s, 1H), 8.00 – 7.95 (m, 2H), 7.79 (d, *J* = 8.8 Hz, 1H), 7.79 – 7.73 (m, 3H), 7.58 – 7.53 (m, 1H), 7.57 – 7.50 (m, 2H), 7.49 (dd, *J* = 2.4, 1.4 Hz, 1H), 7.09 – 7.04 (m, 1H), 6.92 (dd, *J* = 8.6, 2.1 Hz, 1H), 4.86 – 4.78 (m, 1H), 4.26 (dd, *J* = 13.3, 8.7 Hz, 1H), 3.59 (s, 3H), 2.75 (dd, *J* = 16.7, 5.0 Hz, 1H), 2.64 (dd, *J* = 16.4, 8.9 Hz, 1H), 2.28 (t, *J* = 7.5 Hz, 2H), 2.01 – 1.87 (m, 1H), 1.86 – 1.74 (m, 1H). ^{13}C NMR (126 MHz, DMSO) δ 174.3, 172.7, 172.3, 165.8, 158.1, 156.0, 141.5, 138.2, 137.3, 133.8, 129.7, 129.4, 122.9, 121.5, 120.5, 117.7, 114.5, 99.3, 52.6, 52.1, 30.5, 26.6. HRMS (ESI) *m/z* calcd. for $C_{30}H_{27}ClN_4O_9Na^+$ $[M + Na]^+$ 645.1364, found 645.1376.



3.13

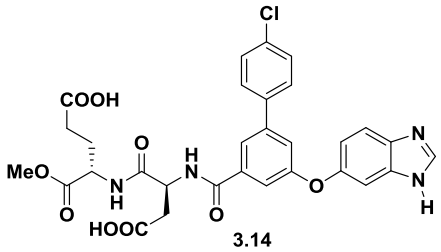
(*S*)-4-((*S*)-3-Carboxy-2-(4'-chloro-5-((3-methyl-1*H*-indazol-6-yl)oxy)-[1,1'-biphenyl]-3-carboxamido)propanamido)-5-methoxy-5-oxopentanoic acid. White solid; yield, 99%; m.p. decomposes at 113.1˚C. $[\alpha]_D^{20}$ -15.6˚ (*c* 5.62, MeOH). ^{1}H NMR (500 MHz, DMSO) δ ppm 12.25 (brs, 2H), 8.84 (d, *J*=8 Hz, 1H), 8.41 (d, *J*=7.5 Hz, 1H), 7.97 (s, 1H), 7.74 (m, 3H), 7.53 (m, 3H), 7.47 (s, 1H), 6.99 (s, 1H), 6.87 (d, *J*=8.5 Hz, 1H), 5.74 (s, 1H), 4.82 (m, 1H), 4.25 (m, 1H), 3.59 (s, 3H), 2.75 (dd, *J*=16.5 Hz, *J*=4.5 Hz, 1H), 2.64 (dd, *J*=16.8 Hz, *J*=9.3 Hz, 1H), 2.46 (s, 3H), 2.28 (m, 2H), 1.95 (m, 1H), 1.79 (m, 1H). ^{13}C NMR (125 MHz, DMSO) δ ppm 174.4, 172.8, 172.4, 471.7, 165.8, 158.3, 156.0, 142.2, 141.9, 141.5, 137.2, 137.2, 133.8, 129.7, 129.4, 122.4, 121.4, 120.4, 119.8, 117.6, 113.6, 99.4, 55.6, 52.6, 52.1, 50.7, 36.7, 30.5, 28.3, 26.5, 12.3. HRMS (ESI) *m/z* calcd. for $C_{31}H_{29}ClN_4O_9Na^+$ $[M + Na]^+$ 659.1521, found 659.1540.



3.31

(*S*)-4-((*S*)-2-(5-((1-Benzyl-1*H*-benzo[*d*]imidazol-6-yl)oxy)-4'-chloro-[1,1'-biphenyl]-3-carboxamido)-3-carboxypropanamido)-5-methoxy-5-oxopentanoic acid (3.31). Off-white solid; yield, 99%; m.p. decomposes at 65˚C. $[\alpha]_D^{20}$ -16.8˚ (*c* 20.0, MeOH). ^{1}H NMR (500 MHz, DMSO-d_6) δ 9.21 (s, 1H), 8.84 (d, *J* = 7.5 Hz, 1H), 8.43 (d,

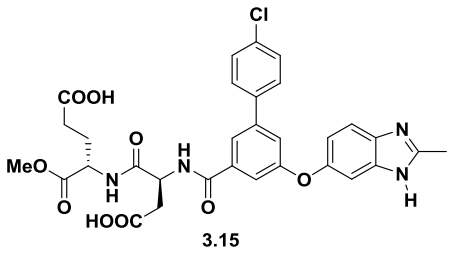
J = 7.6 Hz, 1H), 7.95 (d, *J* = 15.3 Hz, 1H), 7.82 (s, *J* = 8.8 Hz, 1H), 7.71 (d, *J* = 8.1 Hz, 2H), 7.54 (m, 2H), 7.43 (d, *J* = 19.8 Hz, 1H), 7.36 (m, 2H), 7.32 – 7.09 (m, 3H), 5.59 (s, 2H), 4.82 (m, 1H), 4.25 (m, 1H), 3.60 (s, 3H), 2.80 – 2.70 (m, 1H), 2.66 (dd, *J* = 16.7, 9.3 Hz, 1H), 2.27 (m, 2H), 1.97 (m, 1H), 1.81 (m, 1H). ^{13}C NMR (126 MHz, DMSO) δ 174.4, 172.8, 172.3, 171.7, 168.5, 165.8, 158.6, 154.0, 144.3, 141.5, 138.2, 137.2, 135.9, 133.8, 129.7, 129.6, 129.5, 129.5, 129.4, 128.9, 128.9, 128.6, 128.5, 126.0, 121.3, 119.7, 119.2, 117.9, 117.0, 103.6, 52.6, 52.1, 50.8, 49.5, 36.7, 30.5, 26.5. HRMS (ESI) *m/z* calcd. for $C_{37}H_{33}ClN_4O_9Na^+$ $[M + Na]^+$ 735.1834, found 735.1831.



3.14

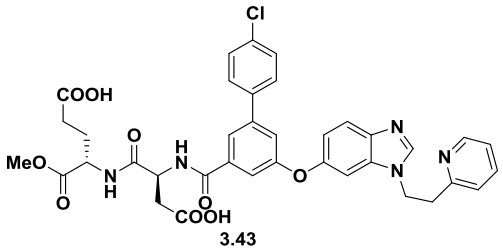
(*S*)-4-((*S*)-2-(5-((1*H*-Benzo[*d*]imidazol-6-yl)oxy)-4'-chloro-[1,1'-biphenyl]-3-carboxamido)-3-carboxypropanamido)-5-methoxy-5-oxopentanoic acid (3.14). White solid; yield, 99%; m.p. decomposes at 123.3˚C. $[\alpha]_D^{20}$ -10.4˚ (*c* 8.94, MeOH). ^{1}H NMR (500 MHz, DMSO-d_6) δ 12.25 (s, 2H), 9.07 (s, 1H), 8.81 (dd, *J* = 12.0, 7.7 Hz, 1H), 8.36 (d, *J* = 7.6 Hz, 1H), 7.97 (t, *J* = 1.5 Hz, 1H), 7.81 (d, *J* = 8.8 Hz, 1H), 7.76 (m, 2H), 7.58 – 7.49 (m, 4H), 7.47 – 7.40 (m, 1H), 7.25 (dd, *J* = 8.8, 2.3 Hz, 1H), 4.87 – 4.76 (m, 1H), 4.25 (td, *J* = 8.2, 5.0 Hz, 1H), 3.59 (s, *3*H), 2.76 (dd, *J* = 16.7, 5.0 Hz, 1H), 2.65 (dd, *J* = 16.6, 9.3 Hz, 1H), 2.30 – 2.19 (m, 2H), 1.95 (dq, *J* = 13.2, 7.2 Hz, 1H), 1.85 – 1.72 (m, 1H). ^{13}C NMR (126 MHz, DMSO) δ 174.3, 174.3, 172.7, 172.3, 171.7, 165.9, 158.6, 154.2, 141.5, 138.2, 137.3, 133.8, 129.7, 129.6, 129.4, 128.9, 121.3, 118.1, 117.0, 105.4, 52.6, 52.1, 50.8, 36.7, 30.5, 26.6, 21.7. HRMS (ESI) *m/z* calcd. for $C_{30}H_{28}ClN_4O_9^-$ $[M + H]^+$ 623.1545,

found 623.1546.



3.15

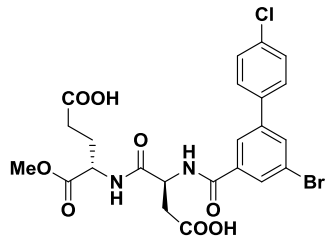
(*S*)-4-((*S*)-3-Carboxy-2-(4'-chloro-5-((2-methyl-1*H*-benzo[*d*]imidazol-6-yl)oxy)-[1,1'-biphenyl]-3-carboxamido)propanamido)-5-methoxy-5-oxopentanoic acid (3.15). White solid; yield, 99%; m.p. decomposes at 133.3˚C. $[\alpha]_D^{20}$ -16.6˚ (*c* 5.82, MeOH). ^{1}H NMR (500 MHz, DMSO-d_6) δ 12.21 (s, 2H) 8.81 (d, *J* = 7.8 Hz, 1H), 8.37 (d, *J* = 7.6 Hz, 1H), 7.94 (s, 1H), 7.77 – 7.73 (m, 2H), 7.71 (d, *J* = 8.9 Hz, 1H), 7.57 – 7.49 (m, 3H), 7.43 – 7.35 (m, 2H), 7.26 – 7.09 (m, 1H), 4.87 – 4.76 (m, 1H), 4.32 – 4.16 (m, 1H), 3.59 (s, 3H), 2.75 (dd, *J* = 16.7, 4.7 Hz, 1H), 2.67 (m, 4H), 2.29 – 2.23 (m, 2H), 2.01 – 1.88 (m, 1H), 1.85 – 1.73 (m, 1H). ^{13}C NMR (126 MHz, DMSO) δ 174.1, 172.5, 172.1, 171.4, 165.6, 158.7, 152.8, 141.2, 138.0, 137.0, 133.6, 129.4, 129.3, 129.1, 128.6, 120.9, 116.5, 52.3, 51.9, 50.5, 36.4, 30.2, 26.3, 13.6. HRMS (ESI) *m/z* calcd. for $C_{31}H_{30}ClN_4O_9^+$ $[M + H]^+$ 637.1701, found 637.1707.



3.43

2-(2-(6-((5-(((*S*)-3-Carboxy-1-(((*S*)-4-carboxy-1-methoxy-1-oxobutan-2-yl)amino)-1-oxopropan-2-yl)carbamoyl)-4'-chloro-[1,1'-biphenyl]-3-yl)oxy)-1*H*-benzo[d]imidazol-1-yl)ethyl)pyridin-1-ium 2,2,2-trifluoroacetate (3.43). Light purple solid; yield, 99%; m.p. decomposes at 60.0˚C. $[\alpha]_D^{20}$ -9.4˚ (*c* 29.22, MeOH). ^{1}H NMR (500

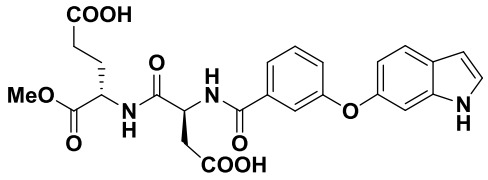
MHz, DMSO-d_6) δ 12.25 (s, 2H), 9.10 (s, 1H), 8.84 (d, *J* = 7.8 Hz, 1H), 8.45 (d, *J* = 5.1 Hz, 1H), 8.39 (d, *J* = 7.5 Hz, 1H), 8.01 – 7.92 (m, 1H), 7.85 – 7.72 (m, 3H), 7.62 – 7.46 (m, 3H), 7.40 (t, *J* = 2.1 Hz, 1H), 7.35 (d, *J* = 7.7 Hz, 2H), 7.30 – 7.08 (m, 4H), 4.87 – 4.80 (m, 1H), 4.78 (t, *J* = 6.8 Hz, 2H), 4.25 (ddd, *J* = 9.3, 7.4, 5.1 Hz, 1H), 3.59 (d, *J* = 4.7 Hz, 3H), 3.38 (t, *J* = 6.8 Hz, 2H), 2.76 (dd, *J* = 16.7, 4.8 Hz, 1H), 2.66 (dd, *J* = 16.6, 9.4 Hz, 1H), 2.31 – 2.21 (m, 4H), 2.00 – 1.89 (m, 1H), 1.86 – 1.73 (m, 1H). ^{13}C NMR (126 MHz, DMSO) δ 174.1, 172.5, 172.1, 171.5, 165.6, 158.8, 158.5, 156.7, 153.9, 148.1, 141.3, 138.0, 137.8, 137.0, 133.6, 133.3, 129.4, 129.3, 129.2, 128.6, 125.7, 124.9, 123.1, 121.0, 119.6, 118.3, 118.2, 116.7, 103.6, 52.3, 51.9, 50.6, 45.4, 36.4, 36.4, 30.2, 26.3, 21.5. HRMS (ESI) *m/z* calcd. for $C_{37}H_{35}ClN_5O_9^+$ $[M + H]^+$ 728.2123, found 728.2130.



3.44

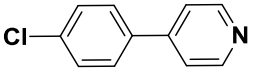
(*S*)-4-((*S*)-2-(5-Bromo-4'-chloro-[1,1'-biphenyl]-3-carboxamido)-3-carboxypropanamido)-5-methoxy-5-oxopentanoic acid (3.44). White solid; yield, 99%; m.p. 187.8—188.3˚C. $[\alpha]_D^{20}$ -12.3˚ (*c* 24.98, MeOH). ^{1}H NMR (500 MHz, DMSO-d_6) δ 12.22 (s, 2H), 8.89 (d, *J* = 7.7 Hz, 1H), 8.40 (d, *J* = 7.6 Hz, 1H), 8.12 (t, *J* = 1.6 Hz, 1H), 8.06 – 7.99 (m, 2H), 7.81 – 7.76 (m, 2H), 7.58 – 7.52 (m, 2H), 4.83 (ddd, *J* = 9.3, 7.8, 4.8 Hz, 1H), 4.34 – 4.23 (m, 1H), 3.61 (s, 3H), 2.77 (dd, *J* = 16.6, 4.7 Hz, 1H), 2.66 (dd, *J* = 16.6, 9.3 Hz, 1H), 2.29 (t, *J* = 7.5 Hz, 2H), 2.02 – 1.91 (m, 1H), 1.82 (ddd, *J* = 13.6, 9.3, 6.9 Hz, 1H). ^{13}C NMR (126 MHz, DMSO) δ 174.1, 172.5, 172.1, 171.3, 165.0, 141.5, 137.2, 137.1, 133.8, 132.3, 129.9, 129.5, 129.3, 125.2, 122.8, 52.3, 51.9, 50.6, 36.4, 30.2,

26.3. HRMS (ESI) *m/z* calcd. for $C_{23}H_{22}BrClN_2O_8Na^+$ [M + Na]$^+$ 591.0146, found 591.0157.



3.45

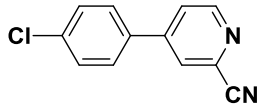
(*S*)-4-((*S*)-2-(3-((1*H*-Indol-6-yl)oxy)benzamido)-3-carboxypropanamido)-5-methoxy-5-oxopentanoic acid (3.45). Light pink solid; yield, 99%; m.p. decomposes at 90˚C. $[\alpha]_D^{20}$ -22.6˚ (*c* 8.54, MeOH). ^{1}H NMR (500 MHz, DMSO-d_6) δ 12.18 (s, 2H), 11.01 (s, 1H), 8.62 (d, *J* = 7.7 Hz, 1H), 8.30 (d, *J* = 7.6 Hz, 1H), 7.57 (d, *J* = 7.8 Hz, 1H), 7.54 (d, *J* = 8.5 Hz, 1H), 7.48 – 7.45 (m, 1H), 7.42 (t, *J* = 7.9 Hz, 1H), 7.30 (t, *J* = 2.7 Hz, 1H), 7.11 (dd, *J* = 8.0, 2.5 Hz, 1H), 7.00 (d, *J* = 2.1 Hz, 1H), 6.75 (dd, *J* = 8.6, 2.2 Hz, 1H), 6.41 (t, *J* = 2.4 Hz, 1H), 4.77 (td, *J* = 8.5, 5.0 Hz, 1H), 4.25 (td, *J* = 8.4, 5.4 Hz, 1H), 3.59 (s, 3H), 2.72 (dd, *J* = 16.7, 4.8 Hz, 1H), 2.65 (dd, *J* = 16.7, 9.2 Hz, 1H), 2.26 (t, *J* = 8.0 Hz, 2H), 1.94 (dq, *J* = 13.4, 7.4 Hz, 1H), 1.85 – 1.73 (m, 1H). ^{13}C NMR (126 MHz, DMSO) δ 174.1, 172.5, 172.1, 171.5, 166.0, 158.7, 151.6, 136.7, 136.2, 130.1, 129.3, 128.6, 126.1, 125.0, 122.0, 121.5, 121.0, 117.2, 112.5, 102.3, 101.6, 52.3, 51.8, 50.5, 36.4, 30.2, 26.4. HRMS (ESI) *m/z* calcd. for $C_{25}H_{26}N_3O_9^+$ [M + H]$^+$ 512.1669, found 512.1670.



3.51

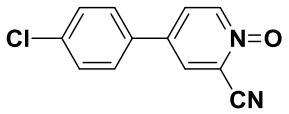
4-(4-Chlorophenyl)pyridine (3.51). To a sealed tube with a magnetic stirrer was added 4-chloropyridine hydrochloride (0.75 g, 5 mmol), (4-chlorophenyl)boronic acid (0.86 g, 5.5 mmol), $Pd(PPh_4)_3$ (0.288 g, 0.25 mmol) and K_2CO_3 (4.3 g, 25 mmol).

Anhydrous dioxane (25 mL) was added. The reaction was heated at 90˚C for 16 h. The reaction mixture was poured into 50 mL saturated $NaHCO_3$ (aq) and extracted with ethyl acetate (50 mL × 2). The organic layers were combined and concentrated *in vacuo*, and the residue was purified by column chromatography (hexanes : ethyl acetate = 1:1) to yield the 0.38 g title compound as a white solid (40%). ^{1}H NMR (300 MHz, Chloroform-*d*) δ 8.70 – 8.62 (m, 2H), 7.62 – 7.53 (m, 2H), 7.49 – 7.43 (m, 3H). ^{13}C NMR (75 MHz, $CDCl_3$) δ 150.6, 136.8, 135.5, 129.6, 129.3, 128.5, 127.9, 127.7, 121.6.



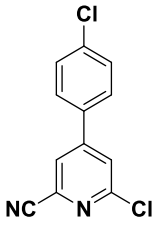
3.53

4-(4-Chlorophenyl)picolinonitrile (3.53). To a stirred solution of 4-(4-chlorophenyl)pyridine 1-oxide (**3.51**) (0.28 g, 1.362 mmol) in 20 mL anhydrous dichloromethane at 0˚, was added TMSCN (204 μL, 1.634 mmol). Dimethylcarbamic chloride (150 μL, 1.634 mmol) was then added dropwise. The reaction was allowed to warm to room temperature and stirred overnight. Upon completion, the reaction was evaporated to dryness under reduced pressure, and the residue was purified by column chromatography (hexanes : ethyl acetate = 10:1) to afford 0.173 g product as light a yellow solid (59%). ^{1}H NMR (500 MHz, Chloroform-*d*) δ 8.75 (d, *J* = 5.2 Hz, 1H), 7.87 (dd, *J* = 1.8, 0.9 Hz, 1H), 7.68 (ddd, *J* = 5.2, 1.9, 0.8 Hz, 1H), 7.60 – 7.55 (m, 2H), 7.53 – 7.48 (m, 2H). ^{13}C NMR (126 MHz, $CDCl_3$) δ 151.8, 148.8, 136.9, 134.9, 134.6, 130.0, 128.5, 126.4, 124.6, 117.4.



3.54

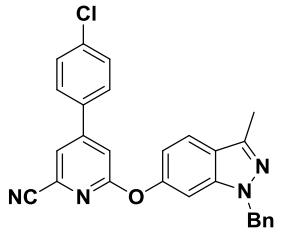
4-(4-Chlorophenyl)-2-cyanopyridine 1-oxide (3.54). To a stirred solution of 4-(4-chlorophenyl)picolinonitrile (**3.53**) (170 mg, 0.792 mmol) in 5 mL chloroform at room temperature, was added mCPBA solution (213 mg, 0.95 mmol in 20 mL chloroform) through an addition funnel over 4 h. The reaction was stirred at room temperature overnight. Upon completion, the reaction mixture was diluted by 20 mL dichloromethane and washed with saturated Na_2CO_3 (aq) (2 ×20 mL). The organic layer was concentrated in vacuo. The residue was purified by flash chromatography (MeOH : DCM = 1:20) to afford 110 mg product as a white solid (60%). ^{1}H NMR (500 MHz, Chloroform-*d*) δ 8.31 (dd, *J* = 7.0, 0.6 Hz, 1H), 7.85 (d, *J* = 2.8 Hz, 1H), 7.64 (dd, *J* = 7.0, 2.7 Hz, 1H), 7.54 – 7.47 (m, 4H). ^{13}C NMR (126 MHz, $CDCl_3$) δ 140.5, 136.7, 136.6, 133.2, 130.2, 128.7, 127.8, 126.7, 111.9.



3.55

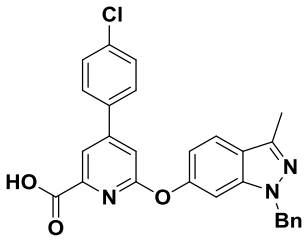
6-Chloro-4-(4-chlorophenyl)picolinonitrile (**3.55**). To a stirred solution of 5 mL $POCl_3$ at room temperature was added 4-(4-chlorophenyl)-2-cyanopyridine 1-oxide (**3.54**) (150 mg, 0.65 mmol). The reaction was heated at 80˚C for 10 h. Upon completion, the reaction mixture was diluted in 20 ethyl acetate and 30 mL saturated $NaHCO_3$ (aq) was added slowly while stirring to quench the unreacted $POCl_3$. The mixture was extracted with 20 mL ethyl acetate. The organic layer was washed with 30 mL brine and concentrated *in*

vacuo. The residue was purified by column chromatography (hexanes : ethyl acetate = 8:2) to afford 102 mg product as a white solid (63%). ^{1}H NMR (500 MHz, Chloroform-*d*) δ 7.80 (s, 1H), 7.71 (d, *J* = 1.8 Hz, 1H), 7.56 (d, *J* = 8.5 Hz, 2H), 7.54 – 7.48 (m, 2H). ^{13}C NMR (126 MHz, $CDCl_3$) δ 153.7, 151.7, 137.6, 134.2, 133.4, 130.2, 128.6, 125.7, 125.4, 116.2.



3.56

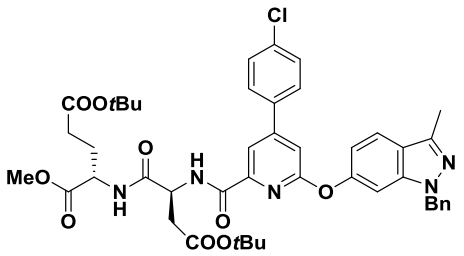
6-((1-Benzyl-3-methyl-1*H*-indazol-6-yl)oxy)-4-(4-chlorophenyl)picolinonitrile (3.56). Procedure from Ullman coupling. White solid, 55%. ^{1}H NMR (500 MHz, Chloroform-*d*) δ 7.67 (dt, *J* = 8.7, 0.9 Hz, 1H), 7.58 (t, *J* = 1.3 Hz, 1H), 7.55 – 7.48 (m, 2H), 7.50 – 7.46 (m, 2H), 7.31 – 7.27 (m, 2H), 7.27 – 7.22 (m, 2H), 7.22 – 7.17 (m, 2H), 7.11 – 7.06 (m, 1H), 6.92 (ddd, *J* = 8.6, 2.0, 1.1 Hz, 1H), 5.49 (s, 2H), 2.60 (d, *J* = 1.1 Hz, 3H). ^{13}C NMR (126 MHz, $CDCl_3$) δ 164.9, 152.3, 152.3, 142.4, 141.2, 137.0, 136.9, 134.5, 131.9, 130.0, 129.0, 128.5, 128.0, 127.3, 122.2, 121.9, 121.9, 116.9, 115.0, 113.2, 101.5, 53.0, 12.2.



3.57

6-((1-Benzyl-3-methyl-1*H*-indazol-6-yl)oxy)-4-(4-chlorophenyl)picolinic acid (3.57). To a stirred solution of 6-((1-benzyl-3-methyl-1*H*-indazol-6-yl)oxy)-4-(4-

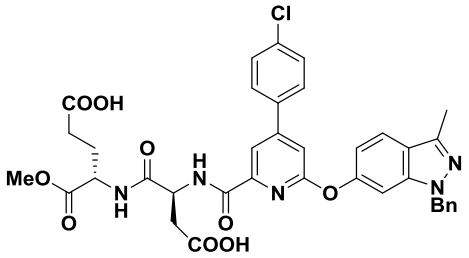
chlorophenyl)picolinonitrile (**3.56**) (30 mg, 0.067 mmol) in MeOH (2 mL) and dioxane (2 mL) was added 6 M NaOH aqueous solution (4 mL). The reaction was heated to reflux for 2 h until no starting material was left. The reaction was cooled to room temperature and neutralized by adding 2 M HCl (aq). The mixture was extracted with ethyl acetate (2 × 20 mL). The organic layers were combined, washed with brine (40 mL) and dried over Na_2SO_4. The solvent was evaporated *in vacuo* to yield 30 mg product, which was used without further purification (96%). ^{1}H NMR (500 MHz, DMSO-d_6) δ 12.84 (s, 2H), 8.06 (d, J = 1.5 Hz, 1H), 7.90 – 7.84 (m, 2H), 7.72 (d, J = 8.6 Hz, 1H), 7.59 – 7.55 (m, 3H), 7.30 – 7.18 (m, 6H), 6.94 (dd, J = 8.7, 1.9 Hz, 1H), 5.49 (s, 2H), 2.47 (s, 3H).



3.58

5-(*tert*-Butyl) 1-methyl ((*S*)-2-(6-((1-benzyl-3-methyl-1*H*-indazol-6-yl)oxy)-4-(4-chlorophenyl)picolinamido)-4-(*tert*-butoxy)-4-oxobutanoyl)-L-glutamate (3.58). General peptide coupling procedure. Colorless syrup, 58%. ^{1}H NMR (500 MHz, Chloroform-*d*) δ 8.54 (d, J = 8.4 Hz, 1H), 8.14 (d, J = 1.5 Hz, 1H), 7.69 (d, J = 8.6 Hz, 1H), 7.56 (d, J = 8.3 Hz, 2H), 7.45 (d, J = 8.5 Hz, 2H), 7.31 – 7.13 (m, 7H), 7.08 (d, J = 1.9 Hz, 1H), 7.00 (dd, J = 8.7, 1.9 Hz, 1H), 5.49 (s, 2H), 4.95 – 4.87 (m, 1H), 4.53 (td, J = 7.7, 5.2 Hz, 1H), 3.71 (s, 3H), 2.88 (dd, J = 17.1, 4.6 Hz, 1H), 2.62 (d, J = 6.7 Hz, 1H), 2.59 (s, 3H), 2.36 – 2.18 (m, 2H), 2.11 (m, 1H), 1.89 (tt, J = 14.6, 7.0 Hz, 1H), 1.37 (s, 9H), 1.34 (s, 9H). ^{13}C NMR (126 MHz, $CDCl_3$) δ 172.1, 172.1, 170.8, 170.4, 164.1, 163.5, 153.0, 152.6, 148.2, 142.3, 141.3, 137.1, 136.3, 135.7, 129.7, 128.9, 128.6, 127.9, 127.4, 121.9,

121.6, 116.0, 114.9, 111.9, 100.9, 99.9, 81.9, 80.9, 52.9, 52.6, 52.1, 49.6, 37.1, 31.4, 29.9, 28.2, 28.1, 12.1.



3.59

(*S*)-4-((*S*)-2-(6-((1-Benzyl-3-methyl-1*H*-indazol-6-yl)oxy)-4-(4-chlorophenyl)picolinamido)-3-carboxypropanamido)-5-methoxy-5-oxopentanoic acid (3.59). General procedure for *tert*-butyl ester deprotection. White solid; yield, 99%; m.p. decomposes at 85.2˚C. $[\alpha]_D^{20}$ -4.8˚ (*c* 6.98, MeOH). ^{1}H NMR (500 MHz, DMSO-d_6) δ 12.15 (s, 2H), 8.55 (d, *J* = 8.4 Hz, 1H), 8.37 (d, *J* = 7.6 Hz, 1H), 8.05 (t, *J* = 1.2 Hz, 1H), 7.77 (dd, *J* = 12.1, 8.5 Hz, 3H), 7.58 – 7.51 (m, 3H), 7.35 (d, *J* = 1.3 Hz, 1H), 7.25 – 7.14 (m, 5H), 7.01 (dt, *J* = 8.3, 1.3 Hz, 1H), 5.50 (s, 2H), 4.80 (q, *J* = 6.7 Hz, 1H), 4.27 (q, *J* = 7.5 Hz, 1H), 3.59 (d, *J* = 1.0 Hz, 3H), 2.71 (d, *J* = 6.2 Hz, 2H), 2.51 (s, 3H, with DMSO), 2.26 (t, *J* = 7.6 Hz, 2H), 1.96 (dt, *J* = 13.0, 6.4 Hz, 1H), 1.79 (dq, *J* = 15.1, 7.8 Hz, 1H). ^{13}C NMR (126 MHz, DMSO) δ 174.1, 172.3, 172.1, 170.7, 163.3, 163.3, 153.3, 151.9, 149.4, 141.7, 141.3, 138.1, 135.5, 135.3, 129.7, 129.4, 128.9, 127.8, 127.8, 122.3, 121.0, 115.6, 114.4, 111.6, 100.8, 52.3, 51.9, 51.8, 49.9, 31.1, 30.2, 26.4, 12.1. HRMS (ESI) *m/z* calcd. for $C_{37}H_{33}ClN_5O_9Na^+$ $[M + Na]^+$ 750.1943, found 750.1961.

3.6 References

1. Hopkins, A. L.; Groom, C. R., The druggable genome. *Nature Reviews. Drug Discovery* **2002,** *1* (9), 727-30.

2. Geschwindner, S.; Ulander, J.; Johansson, P., Ligand binding thermodynamics in drug discovery: still a hot tip? *Journal of Medicinal Chemistry* **2015,** *58* (16), 6321-35.

3. Freire, E., A thermodynamic approach to the affinity optimization of drug candidates. *Chemical Biology & Drug Design* **2009,** *74* (5), 468-72.

4. Holdgate, G. A., Thermodynamics of binding interactions in the rational drug design process. *Expert Opinion on Drug Discovery* **2007,** *2* (8), 1103-14.

5. Ferenczy, G. G.; Keseru, G. M., Enthalpic efficiency of ligand binding. *Journal of Chemical Information and Modeling* **2010,** *50* (9), 1536-41.

6. Garbett, N. C.; Chaires, J. B., Thermodynamic studies for drug design and screening. *Expert Opinion on Drug Discovery* **2012,** *7* (4), 299-314.

7. Klebe, G., Applying thermodynamic profiling in lead finding and optimization. *Nature Reviews. Drug Discovery* **2015,** *14* (2), 95-110.

8. Chang, C. E.; Chen, W.; Gilson, M. K., Ligand configurational entropy and protein binding. *Proceedings of the National Academy of Sciences of the United States of America* **2007,** *104* (5), 1534-9.

9. Freire, E., Do enthalpy and entropy distinguish first in class from best in class? *Drug Discovery Today* **2008,** *13* (19-20), 869-74.

10. Olsson, T. S.; Williams, M. A.; Pitt, W. R.; Ladbury, J. E., The thermodynamics of protein-ligand interaction and solvation: insights for ligand design. *Journal of Molecular Biology* **2008,** *384* (4), 1002-17.

11. Ladbury, J. E.; Klebe, G.; Freire, E., Adding calorimetric data to decision making in lead discovery: a hot tip. *Nature Reviews. Drug Discovery* **2010,** *9* (1), 23-7.

12. Kawasaki, Y.; Freire, E., Finding a better path to drug selectivity. *Drug Discovery today* **2011,** *16* (21-22), 985-90.

13. Udugamasooriya, D. G.; Spaller, M. R., Conformational constraint in protein ligand design and the inconsistency of binding entropy. *Biopolymers* **2008,** *89* (8), 653-667.

14. Kessler, H., Conformation and biological activity of cyclic peptides. *Angewandte Chemie International Edition in English* **1982,** *21* (7), 512-523.

15. Toniolo, C., Conformationally restricted peptides through short-range cyclizations. *International Journal of Peptide and Protein Research* **1990,** *35* (4), 287-300.

16. Rizo, J.; Gierasch, L. M., Constrained peptides: models of bioactive peptides and protein substructures. *Annual Review of Biochemistry* **1992,** *61*, 387-418.

17. Hruby, V. J., Designing peptide receptor agonists and antagonists. *Nature Reviews. Drug Discovery* **2002,** *1* (11), 847-58.

18. Ripka, A. S.; Rich, D. H., Peptidomimetic design. *Current Opinion in Chemical Biology* **1998,** *2* (4), 441-52.

19. DiMaio, J.; Nguyen, T. M.; Lemieux, C.; Schiller, P. W., Synthesis and pharmacological characterization in vitro of cyclic enkephalin analogues: effect of conformational constraints on opiate receptor selectivity. *Journal of Medicinal Chemistry* **1982,** *25* (12), 1432-8.

20. Shimamoto, K.; Ohfune, Y., Syntheses and conformational analyses of glutamate analogs: 2-(2-carboxy-3-substituted-cyclopropyl)glycines as useful probes for excitatory amino acid receptors. *Journal of Medicinal Chemistry* **1996,** *39* (2), 407-23.

21. Liao, S.; Alfaro-Lopez, J.; Shenderovich, M. D.; Hosohata, K.; Lin, J.; Li, X.; Stropova, D.; Davis, P.; Jernigan, K. A.; Porreca, F.; Yamamura, H. I.; Hruby, V. J., De novo design, synthesis, and biological activities of high-affinity and selective non-peptide agonists of the delta-opioid receptor. *Journal of Medicinal Chemistry* **1998,** *41* (24), 4767-76.

22. Suich, D. J.; Mousa, S. A.; Singh, G.; Liapakis, G.; Reisine, T.; DeGrado, W. F., Template-constrained cyclic peptide analogues of somatostatin: subtype-selective binding to somatostatin receptors and antiangiogenic activity. *Bioorganic & Medicinal Chemistry* **2000,** *8* (9), 2229-41.

23. Ying, J.; Gu, X.; Cai, M.; Dedek, M.; Vagner, J.; Trivedi, D. B.; Hruby, V. J., Design, synthesis, and biological evaluation of new cyclic melanotropin peptide analogues selective for the human melanocortin-4 receptor. *Journal of Medicinal Chemistry* **2006,** *49* (23), 6888-96.

24. Tyndall, J. D.; Reid, R. C.; Tyssen, D. P.; Jardine, D. K.; Todd, B.; Passmore, M.; March, D. R.; Pattenden, L. K.; Bergman, D. A.; Alewood, D.; Hu, S. H.; Alewood, P. F.; Birch, C. J.; Martin, J. L.; Fairlie, D. P., Synthesis, stability, antiviral activity, and protease-bound structures of substrate-mimicking constrained macrocyclic inhibitors of HIV-1 protease. *Journal of Medicinal Chemistry* **2000,** *43* (19), 3495-504.

25. Gosselin, F.; Tourwe, D.; Ceusters, M.; Meert, T.; Heylen, L.; Jurzak, M.; Lubell, W. D., Probing opioid receptor-ligand interactions by employment of indolizidin-9-one amino acid as a constrained Gly(2)-Gly(3) surrogate in a leucine-enkephalin

mimic. *The Journal of Peptide Research : Official Journal of the American Peptide Society* **2001,** *57* (4), 337-44.

26. Witty, D. R.; Bateson, J.; Hervieu, G. J.; Al-Barazanji, K.; Jeffrey, P.; Hamprecht, D.; Haynes, A.; Johnson, C. N.; Muir, A. I.; O'Hanlon, P. J.; Stemp, G.; Stevens, A. J.; Thewlis, K.; Winborn, K. Y., Discovery of potent and stable conformationally constrained analogues of the MCH R1 antagonist SB-568849. *Bioorganic & Medicinal Chemistry Letters* **2006,** *16* (18), 4872-8.

27. Szewczuk, Z.; Gibbs, B. F.; Yue, S. Y.; Purisima, E. O.; Konishi, Y., Conformationally restricted thrombin inhibitors resistant to proteolytic digestion. *Biochemistry* **1992,** *31* (38), 9132-40.

28. Satoh, T.; Aramini, J. M.; Li, S.; Friedman, T. M.; Gao, J.; Edling, A. E.; Townsend, R.; Koch, U.; Choksi, S.; Germann, M. W.; Korngold, R.; Huang, Z., Bioactive peptide design based on protein surface epitopes. A cyclic heptapeptide mimics CD4 domain 1 CC' loop and inhibits CD4 biological function. *The Journal of Biological Chemistry* **1997,** *272* (18), 12175-80.

29. Schmiedeberg, N.; Schmitt, M.; Rolz, C.; Truffault, V.; Sukopp, M.; Burgle, M.; Wilhelm, O. G.; Schmalix, W.; Magdolen, V.; Kessler, H., Synthesis, solution structure, and biological evaluation of urokinase type plasminogen activator (uPA)-derived receptor binding domain mimetics. *Journal of Medicinal Chemistry* **2002,** *45* (23), 4984-94.

30. Rozek, A.; Powers, J. P.; Friedrich, C. L.; Hancock, R. E., Structure-based design of an indolicidin peptide analogue with increased protease stability. *Biochemistry* **2003,** *42* (48), 14130-8.

31. Pakkala, M.; Hekim, C.; Soininen, P.; Leinonen, J.; Koistinen, H.; Weisell, J.; Stenman, U. H.; Vepsalainen, J.; Narvanen, A., Activity and stability of human kallikrein-2-specific linear and cyclic peptide inhibitors. *Journal of Peptide Science : an Official Publication of the European Peptide Society* **2007,** *13* (5), 348-53.

32. Matsuzaki, K.; Yoneyama, S.; Fujii, N.; Miyajima, K.; Yamada, K.; Kirino, Y.; Anzai, K., Membrane permeabilization mechanisms of a cyclic antimicrobial peptide, tachyplesin I, and its linear analog. *Biochemistry* **1997,** *36* (32), 9799-806.

33. Gangwar, S.; Jois, S. D.; Siahaan, T. J.; Vander Velde, D. G.; Stella, V. J.; Borchardt, R. T., The effect of conformation on membrane permeability of an acyloxyalkoxy-linked cyclic prodrug of a model hexapeptide. *Pharmaceutical Research* **1996,** *13* (11), 1657-62.

34. Tamura, K.; Agrios, K. A.; Vander Velde, D.; Aube, J.; Borchardt, R. T., Effect of stereochemistry on the transport of Aca-linked beta-turn peptidomimetics across a human intestinal cell line. *Bioorganic & Medicinal Chemistry* **1997,** *5* (9), 1859-66.

35. Gudmundsson, O. S.; Jois, S. D.; Vander Velde, D. G.; Siahaan, T. J.; Wang, B.;

Borchardt, R. T., The effect of conformation on the membrane permeation of coumarinic acid- and phenylpropionic acid-based cyclic prodrugs of opioid peptides. *The Journal of Peptide Research : Official Journal of the American Peptide Society* **1999,** *53* (4), 383-92.

36. Rezai, T.; Bock, J. E.; Zhou, M. V.; Kalyanaraman, C.; Lokey, R. S.; Jacobson, M. P., Conformational flexibility, internal hydrogen bonding, and passive membrane permeability: successful in silico prediction of the relative permeabilities of cyclic peptides. *Journal of the American Chemical Society* **2006,** *128* (43), 14073-80.

37. Gonzalez-Ruiz, D.; Gohlke, H., Targeting protein-protein interactions with small molecules: challenges and perspectives for computational binding epitope detection and ligand finding. *Current Medicinal Chemistry* **2006,** *13* (22), 2607-25.

38. Sundberg, E. J.; Mariuzza, R. A., Luxury accommodations: the expanding role of structural plasticity in protein-protein interactions. *Structure (London, England : 1993)* **2000,** *8* (7), R137-42.

39. Feher, V. A.; Cavanagh, J., Millisecond-timescale motions contribute to the function of the bacterial response regulator protein Spo0F. *Nature* **1999,** *400* (6741), 289-93.

40. Kay, L. E.; Muhandiram, D. R.; Farrow, N. A.; Aubin, Y.; Forman-Kay, J. D., Correlation between dynamics and high affinity binding in an SH2 domain interaction. *Biochemistry* **1996,** *35* (2), 361-8.

41. Bista, M.; Wolf, S.; Khoury, K.; Kowalska, K.; Huang, Y.; Wrona, E.; Arciniega, M.; Popowicz, G. M.; Holak, T. A.; Domling, A., Transient protein states in designing inhibitors of the MDM2-p53 interaction. *Structure (London, England : 1993)* **2013,** *21* (12), 2143-51.

42. Wells, J. A.; McClendon, C. L., Reaching for high-hanging fruit in drug discovery at protein-protein interfaces. *Nature* **2007,** *450* (7172), 1001-9.

43. Popowicz, G. M.; Czarna, A.; Wolf, S.; Wang, K.; Wang, W.; Domling, A.; Holak, T. A., Structures of low molecular weight inhibitors bound to MDMX and MDM2 reveal new approaches for p53-MDMX/MDM2 antagonist drug discovery. *Cell Cycle (Georgetown, Tex.)* **2010,** *9* (6), 1104-11.

44. Huang, Z.; Zhang, M.; Burton, S. D.; Katsakhyan, L. N.; Ji, H., Targeting the Tcf4 G13ANDE17 binding site to selectively disrupt beta-catenin/T-cell factor protein-protein interactions. *ACS Chemical Biology* **2014,** *9* (1), 193-201.

45. Morris, G. M.; Huey, R.; Lindstrom, W.; Sanner, M. F.; Belew, R. K.; Goodsell, D. S.; Olson, A. J., AutoDock4 and AutoDockTools4: automated docking with selective receptor flexibility. *Journal of Computational Chemistry* **2009,** *30* (16), 2785-91.

46. Wei, Z.-L.; Gowlugari, S.; Kaub, C.; Wang, Z.; Cao, Y.; Kincaid, J. Preparation of biarylamide compounds and their aza analogs as purinergic P2X2 and P2X3 modulators useful in the treatment of pain and other diseases. WO2010033168A2, 2010.

47. Ma, D.; Cai, Q., N,N-Dimethyl glycine-promoted Ullmann coupling reaction of phenols and aryl halides. *Organic Letter* **2003,** *5* (21), 3799-3802.

48. Zhang, M.; Huang, Z.; Yu, B.; Ji, H., New homogeneous high-throughput assays for inhibitors of beta-catenin/Tcf protein-protein interactions. *Analytical Biochemistry* **2012,** *424* (1), 57-63.

49. Yu, B.; Huang, Z.; Zhang, M.; Dillard, D. R.; Ji, H., Rational design of small-molecule inhibitors for beta-catenin/T-cell factor protein-protein interactions by bioisostere replacement. *ACS Chemical Biology* **2013,** *8* (3), 524-9.

50. Wada, Y.; Shirahashi, H.; Iwanami, T.; Ogawa, M.; Nakano, S.; Morimoto, A.; Kasahara, K.-i.; Tanaka, E.; Takada, Y.; Ohashi, S.; Mori, M.; Shuto, S., Discovery of novel indazole derivatives as highly potent and selective human β3-adrenergic receptor agonists with the possibility of having no cardiovascular side effects. *Journal of Medicinal Chemistry* **2015,** *58* (15), 6048-6057.

51. Hanan, E. J.; Chan, B. K.; Estrada, A. A.; Shore, D. G.; Lyssikatos, J. P., Mild and general one-pot reduction and cyclization of aromatic and heteroaromatic 2-nitroamines to bicyclic 2H-imidazoles. *Synlett* **2010,** (18), 2759-2764.

52. Haddach, A. A.; Kelleman, A.; Deaton-Rewolinski, M. V., An efficient method for the N-debenzylation of aromatic heterocycles. *Tetrahedron Letters* **2002,** *43* (3), 399-402.

53. Zhang, M.; Wisniewski, J. A.; Ji, H., AlphaScreen selectivity assay for β-catenin/B-cell lymphoma 9 inhibitors. *Analytical Biochemistry* **2015,** *469*, 43-53.

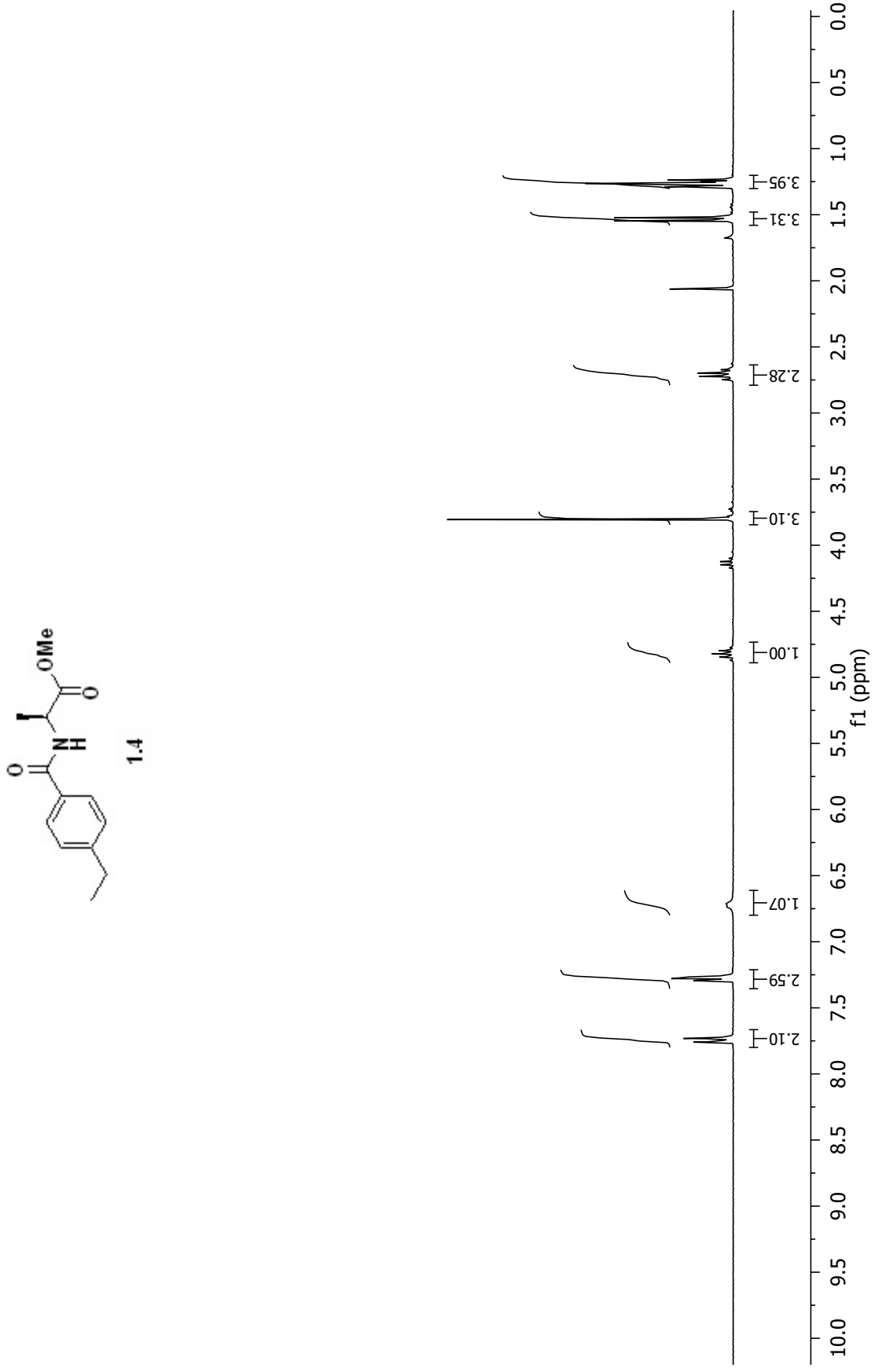
54. Sampietro, J.; Dahlberg, C. L.; Cho, U. S.; Hinds, T. R.; Kimelman, D.; Xu, W., Crystal structure of a beta-catenin/BCL9/Tcf4 complex. *Molecular Cell* **2006,** *24* (2), 293-300.

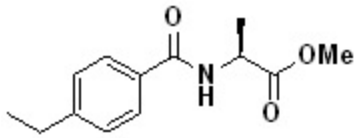
55. Zhang, M.; Catrow, J. L.; Ji, H., High-throughput selectivity assays for small-molecule inhibitors of beta-catenin/T-cell factor protein-protein interactions. *ACS Medicinal Chemistry Letters* **2013,** *4* (2), 306-11.

56. Nikolovska-Coleska, Z.; Wang, R.; Fang, X.; Pan, H.; Tomita, Y.; Li, P.; Roller, P. P.; Krajewski, K.; Saito, N. G.; Stuckey, J. A.; Wang, S., Development and optimization of a binding assay for the XIAP BIR3 domain using fluorescence polarization. *Analytical Biochemistry* **2004,** *332* (2), 261-73.

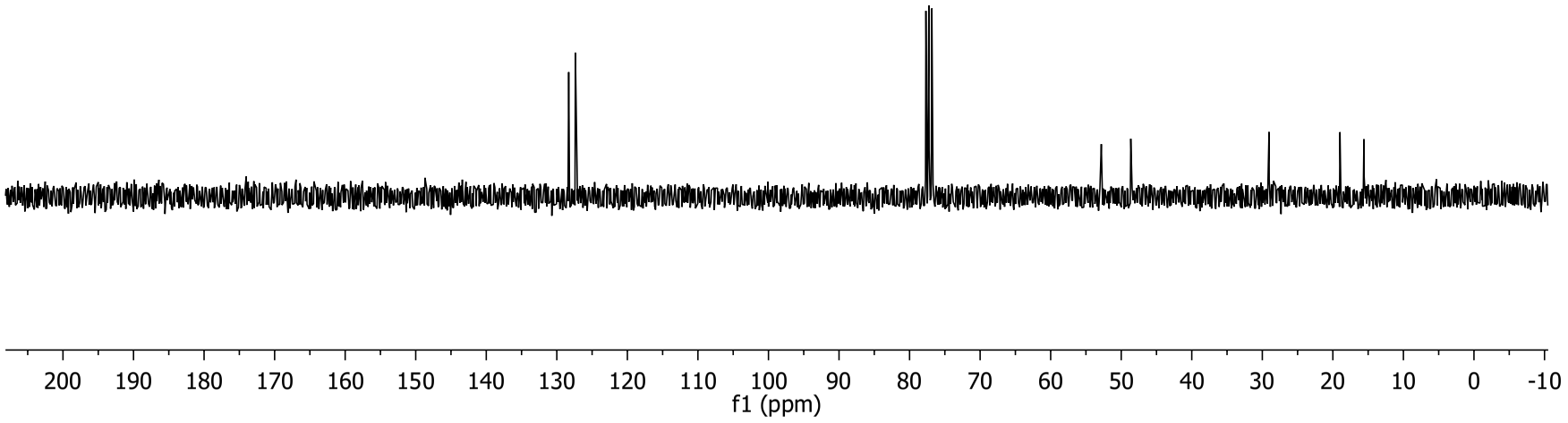
APPENDIX A

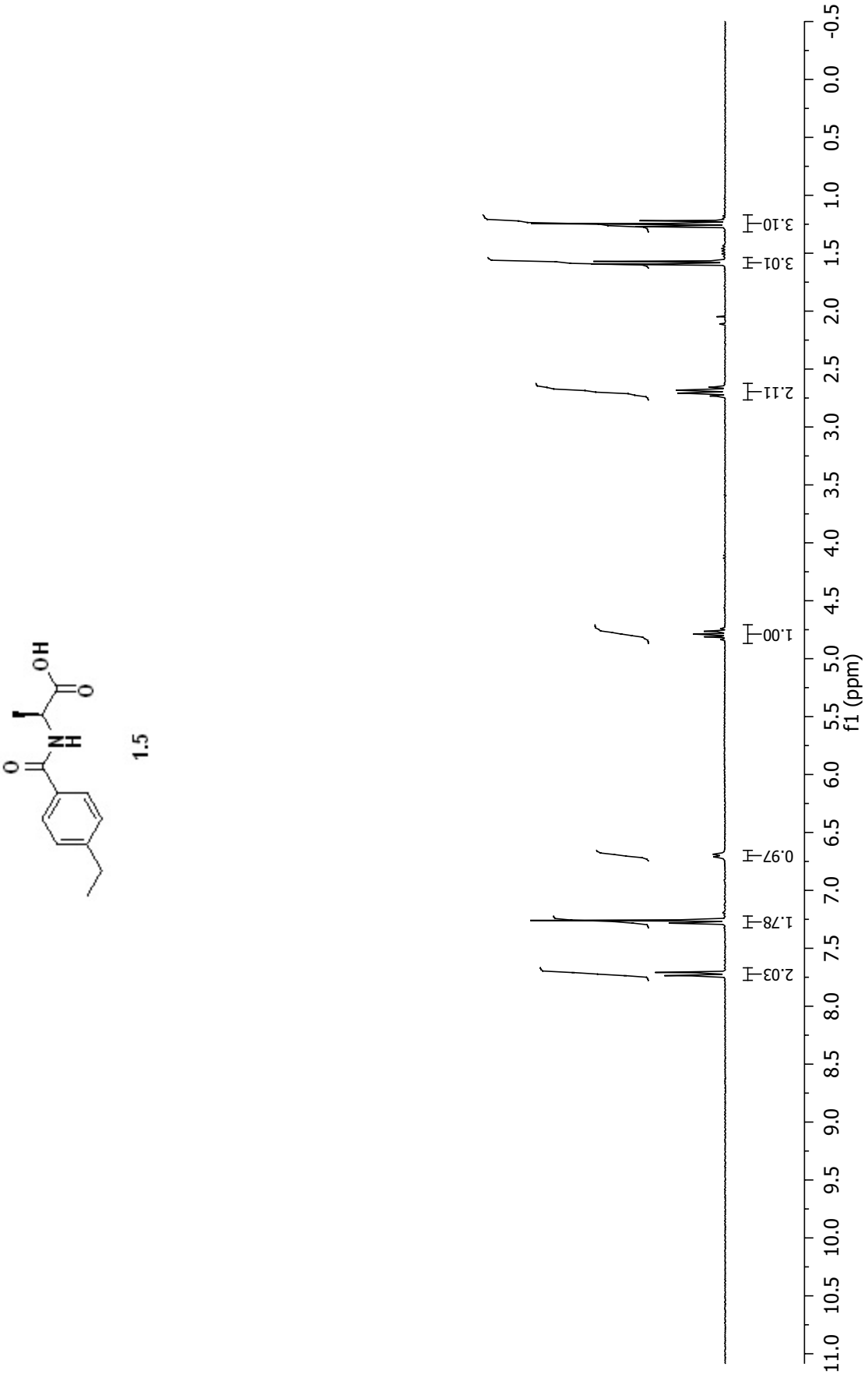
^{1}H AND ^{13}C NMR SPECTRA OF CHAPTER 1

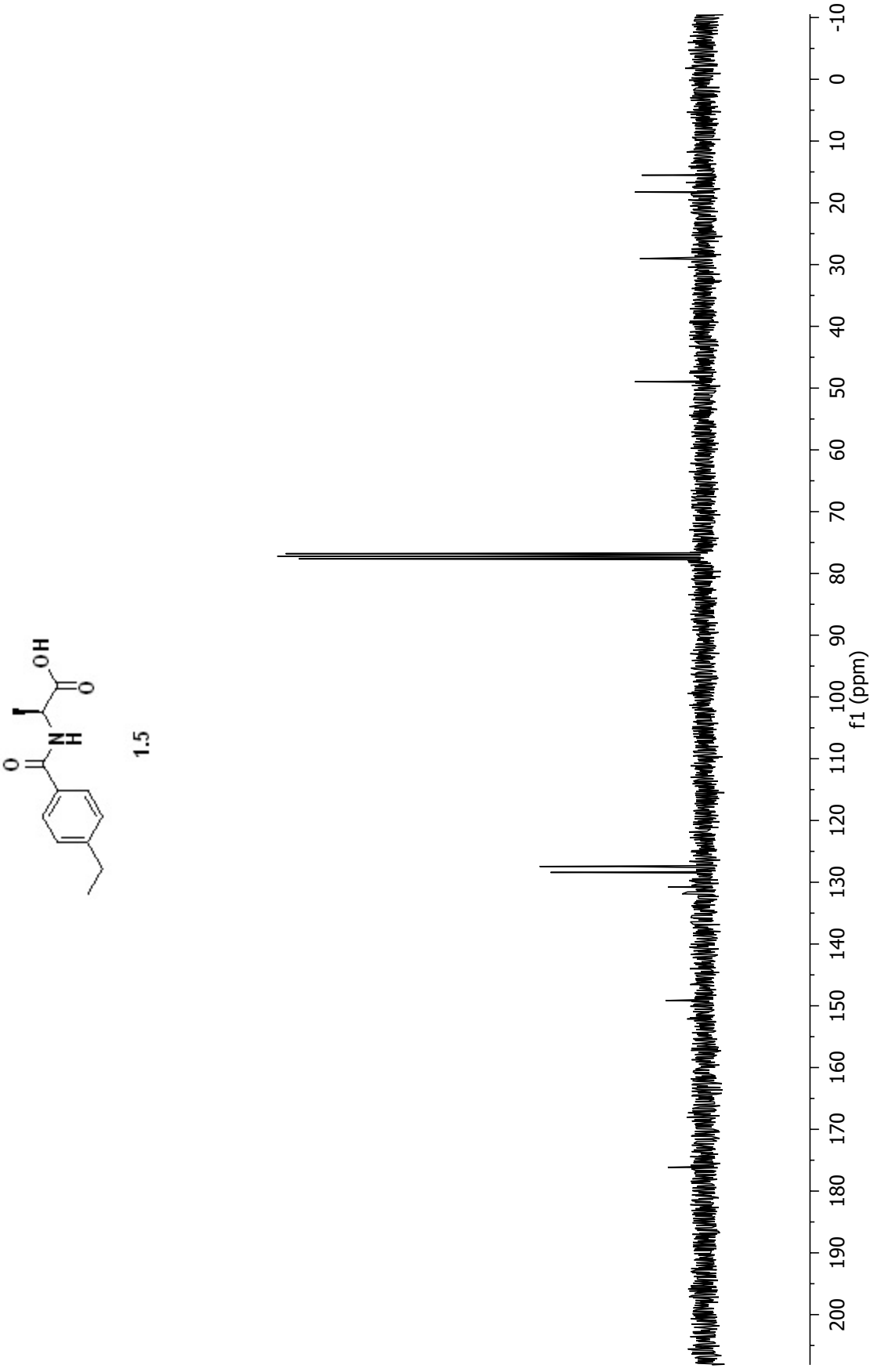

OMe
O
H
N
O
1.4
2.10
2.59
1.07
1.00
3.10
2.28
3.31
3.95
10.0
9.5
9.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
f1 (ppm)

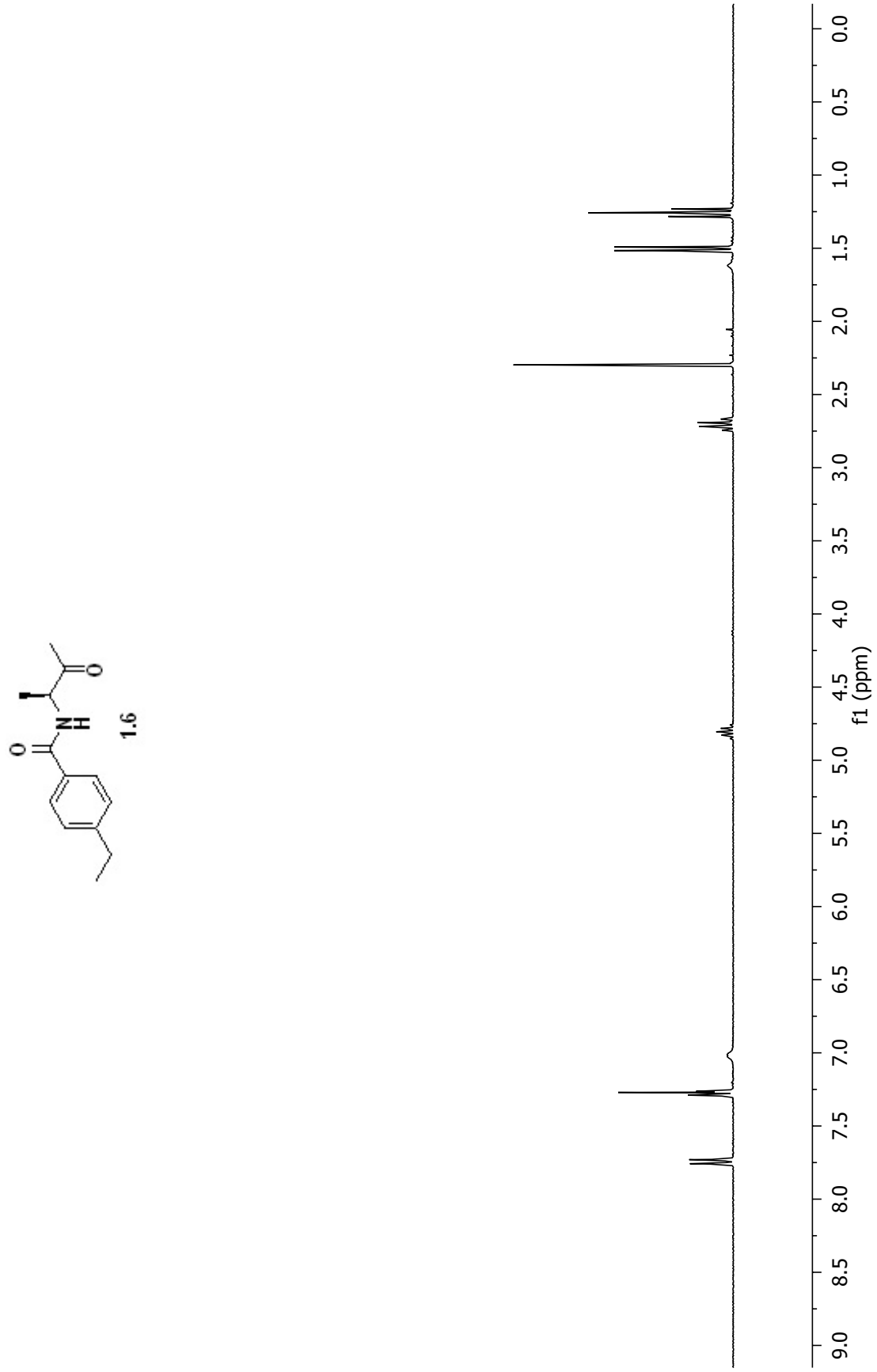

O
N
H
OMe
O

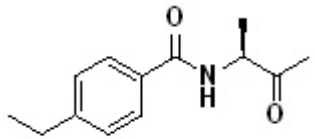
1.4


200
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
-10
f1 (ppm)

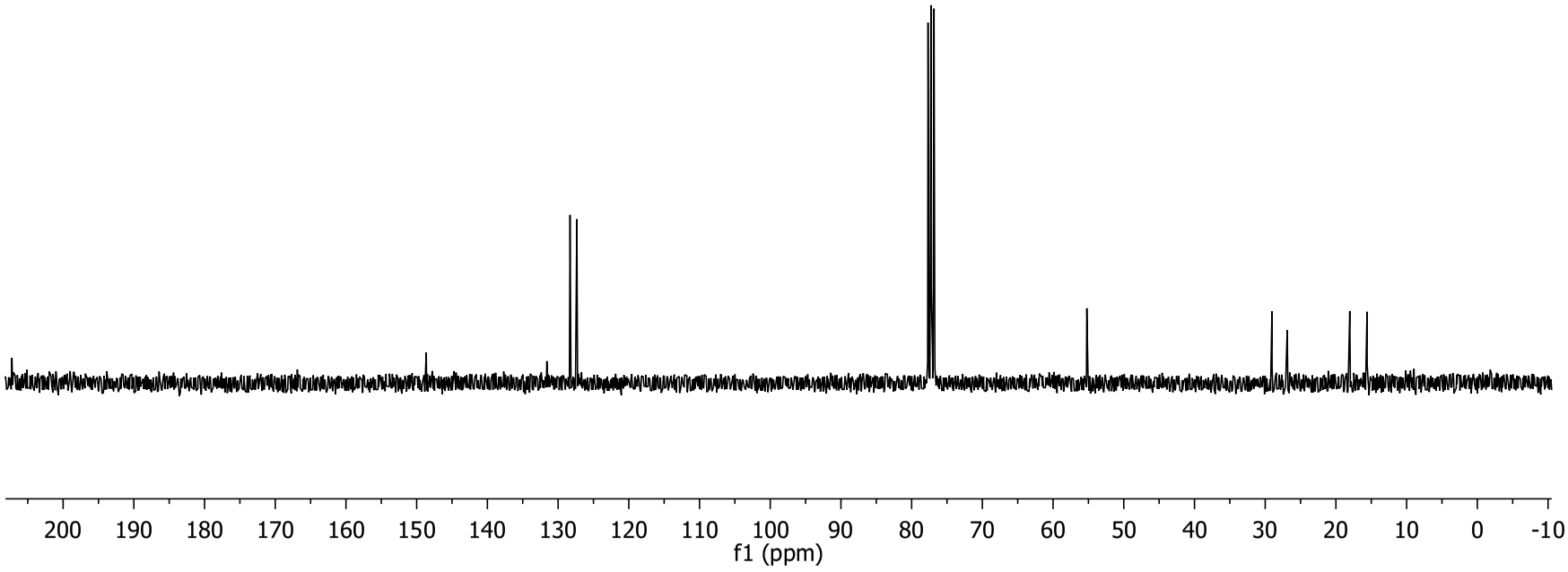

1.5
3.10
3.01
2.11
1.00
0.97
1.78
2.03
f1 (ppm)

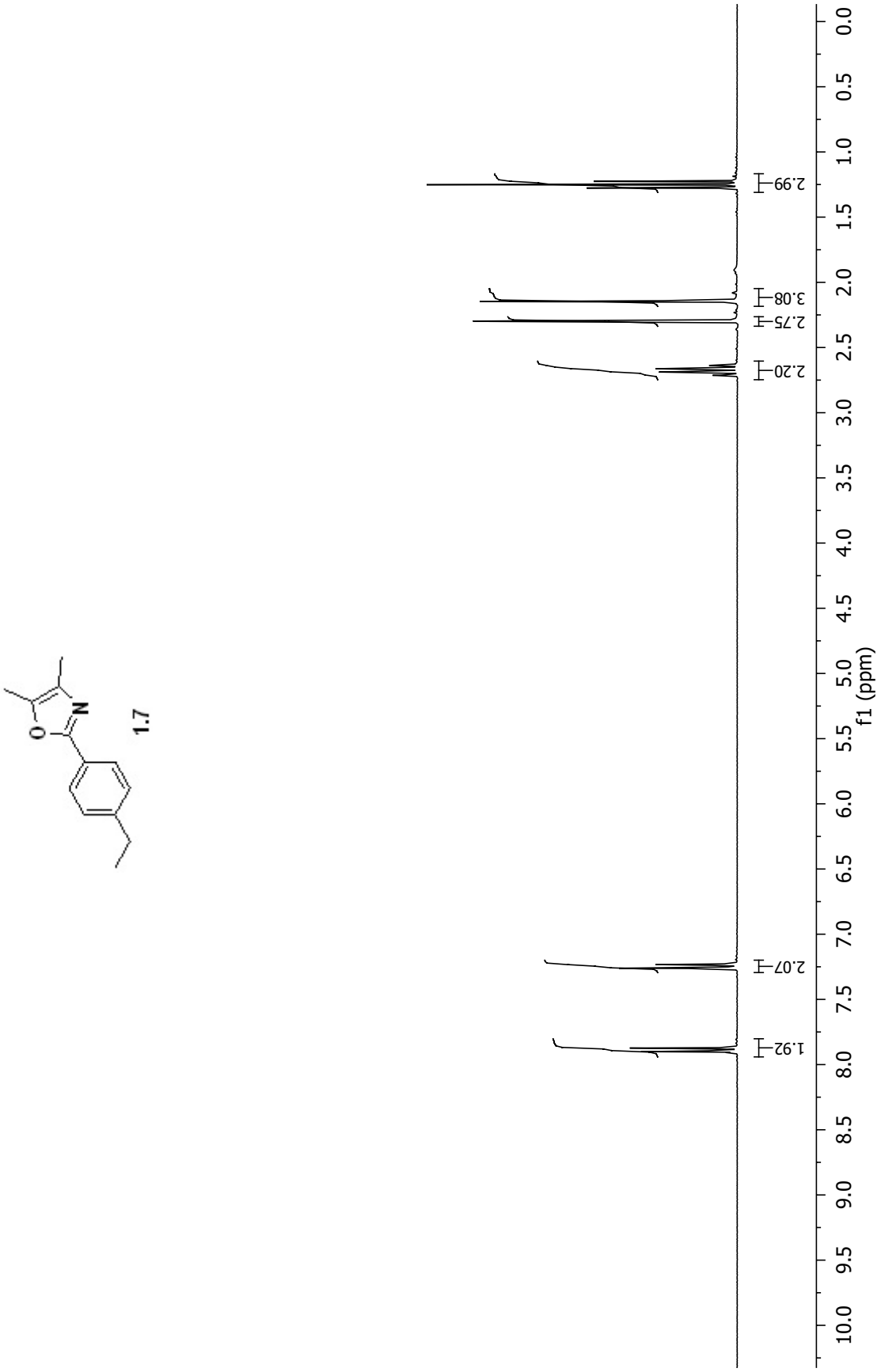

OH
O
N
H
O
1.5
-10
0
10
20
30
40
50
60
70
80
90
100
110
120
130
140
150
160
170
180
190
200
f1 (ppm)


O
N
H
O
1.6
9.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
f1 (ppm)

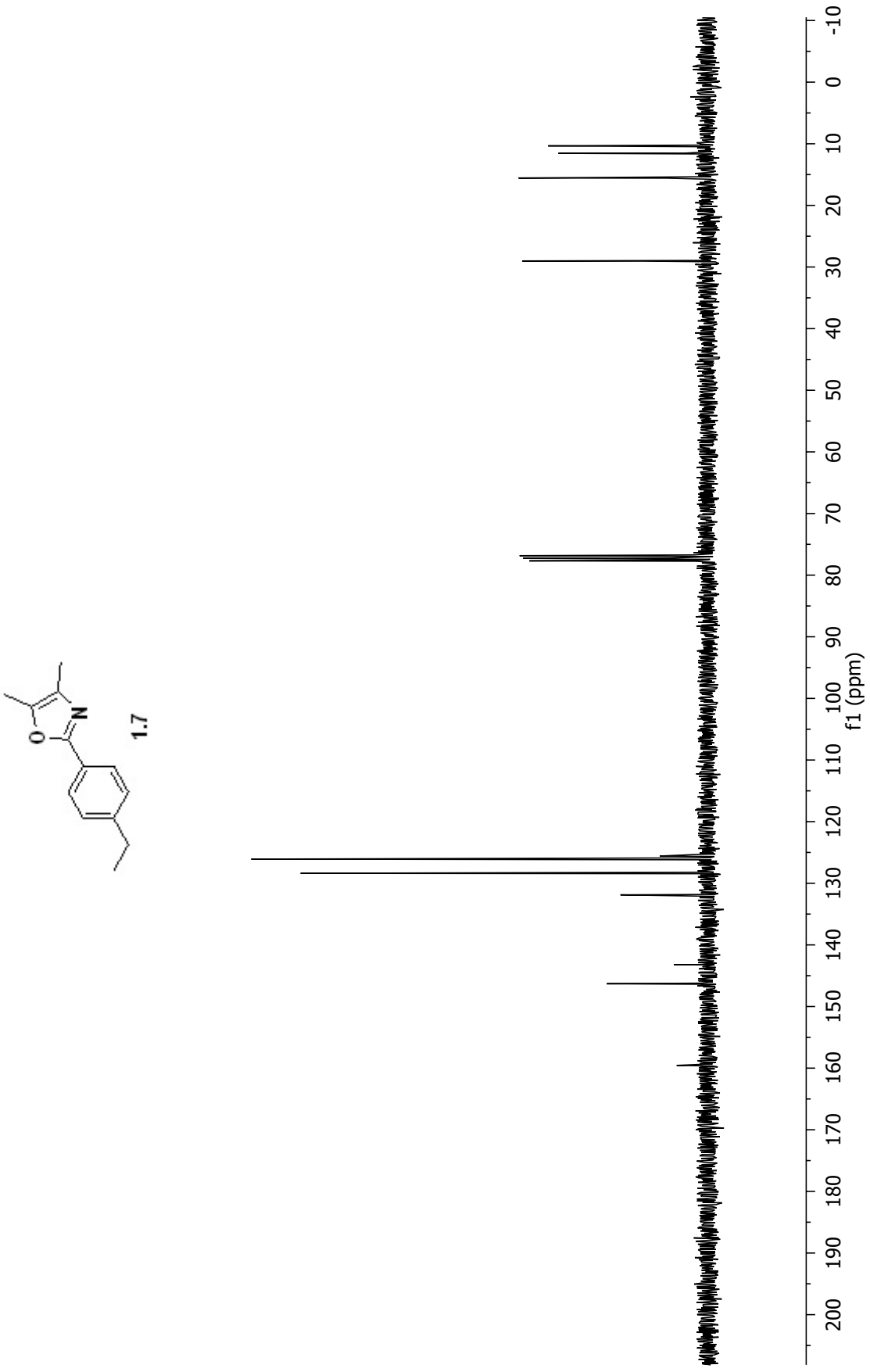

O
N
H
O

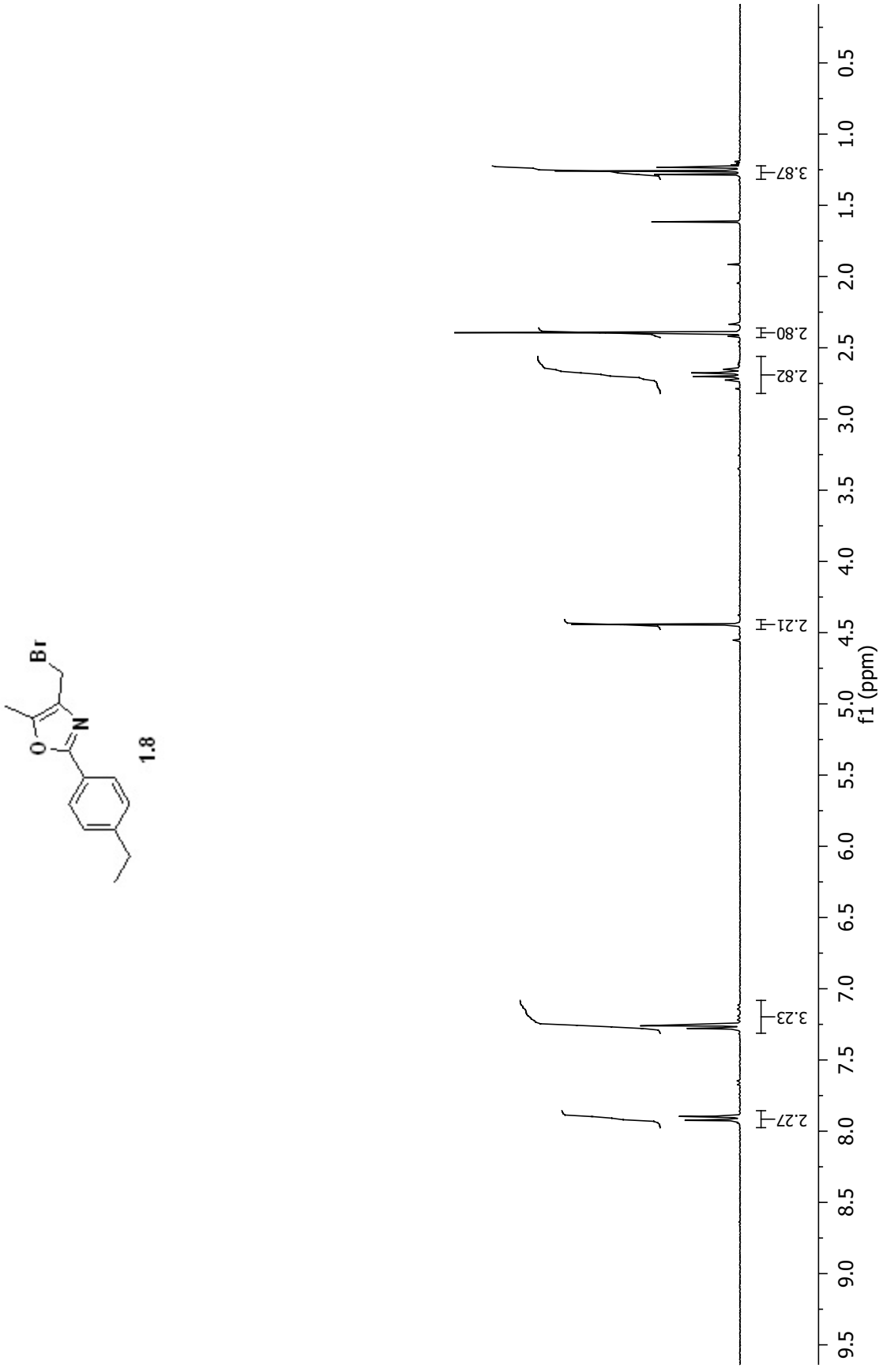
1.6

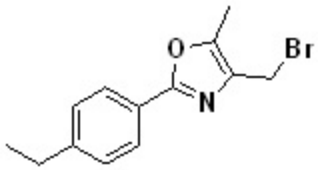

200
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
-10
f1 (ppm)


1.92
2.07
2.20
2.75
3.08
2.99
f1 (ppm)
10.0
9.5
9.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
O
N

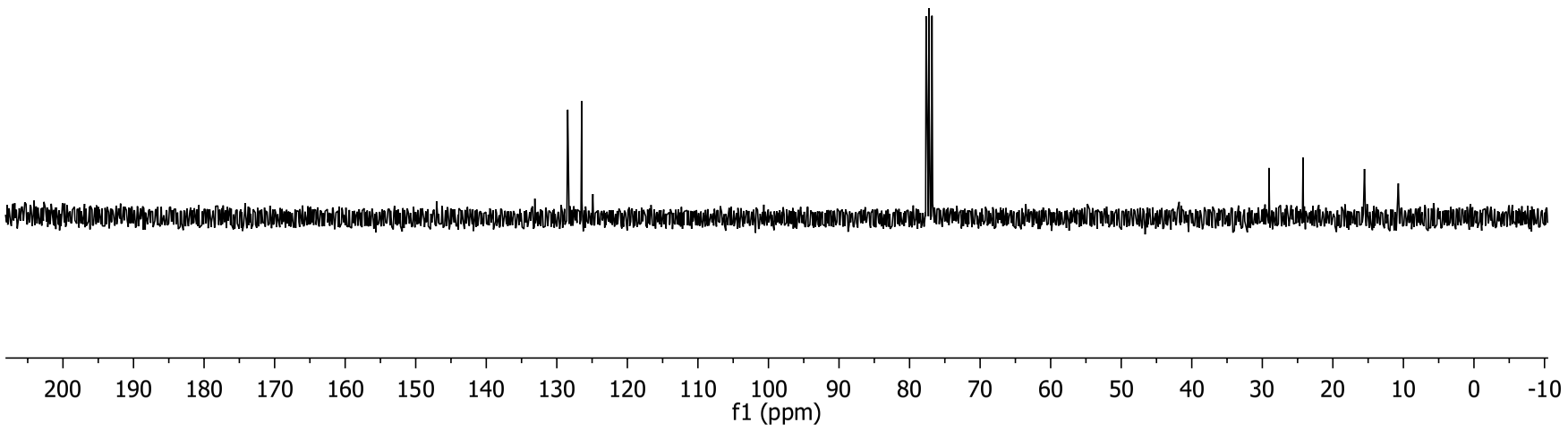
1.7

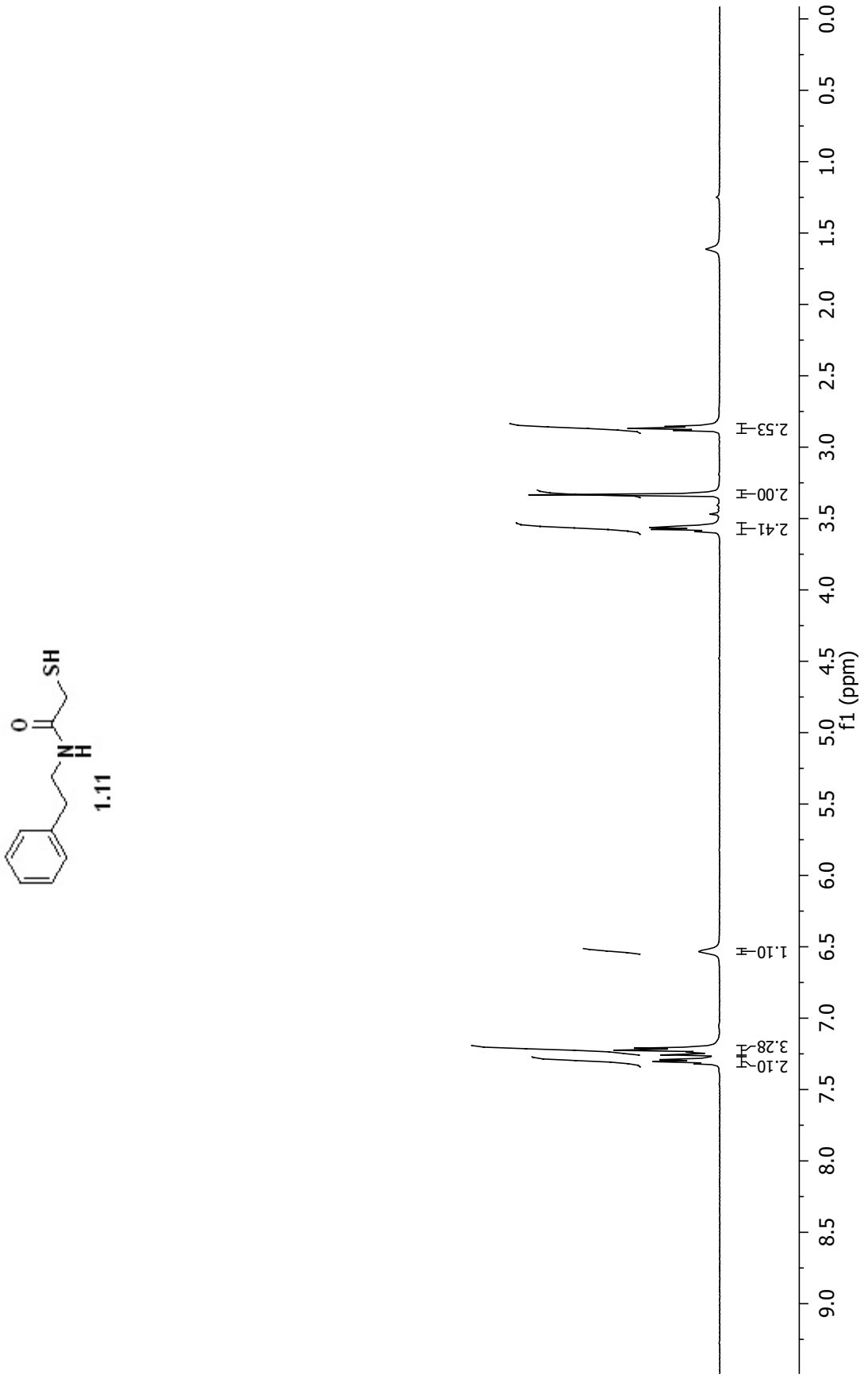

1.7
f1 (ppm)

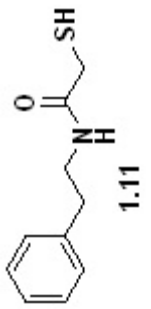

Br
O
N
1.8
2.27
3.23
2.21
2.82
2.80
3.87
9.5
9.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
f1 (ppm)


O
N
Br

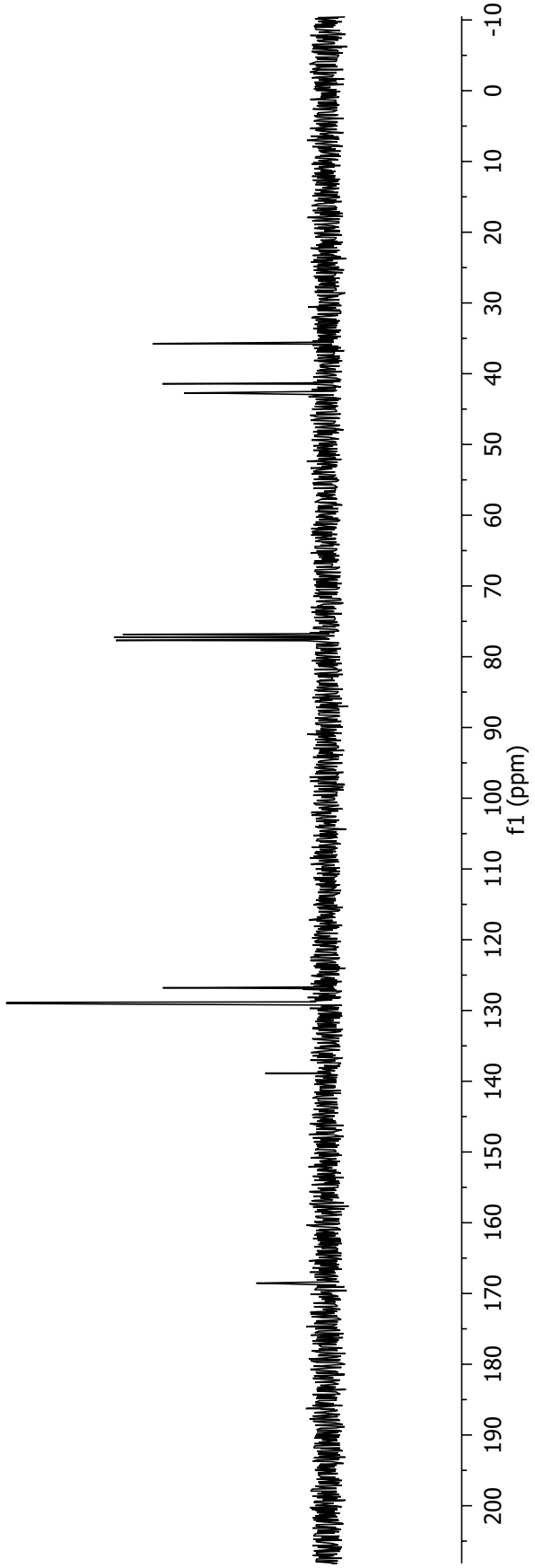
1.8

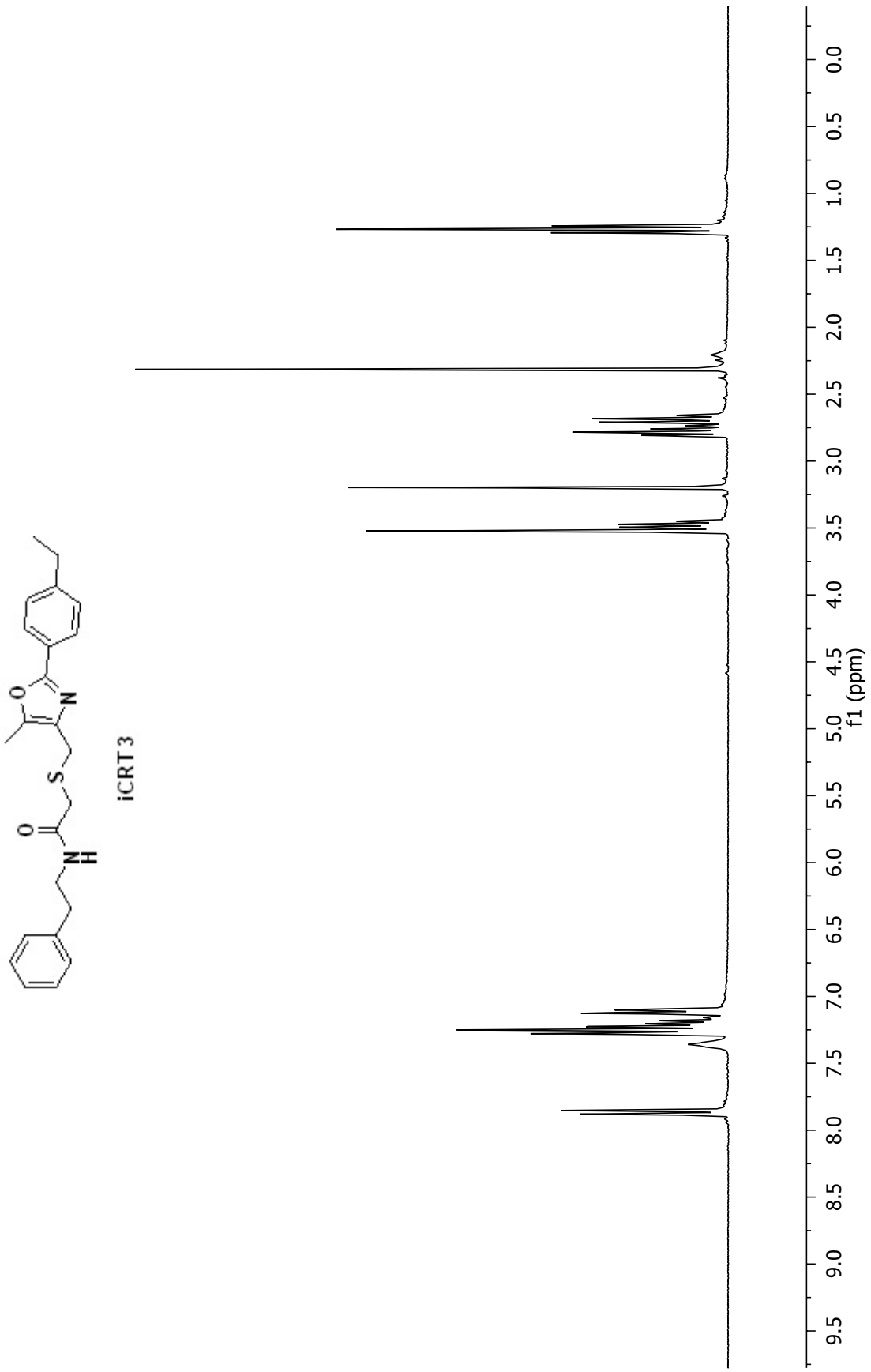

200
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
-10
f1 (ppm)

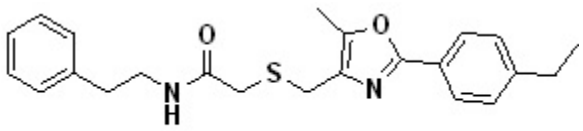

SH
O
N
H
1.11
2.53
2.00
2.41
1.10
3.28
2.10
0.0
0.5
1.0
1.5
2.0
2.5
3.0
3.5
4.0
4.5
5.0
5.5
6.0
6.5
7.0
7.5
8.0
8.5
9.0
f1 (ppm)


SH
O
N
H

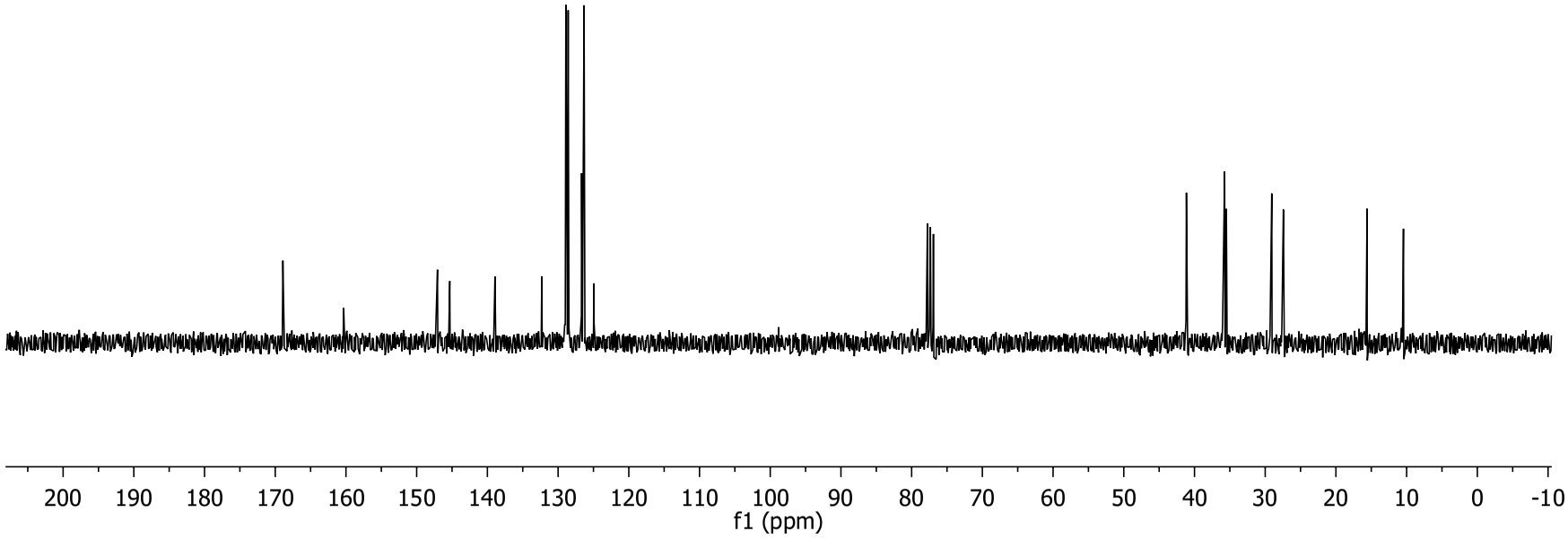
1.11

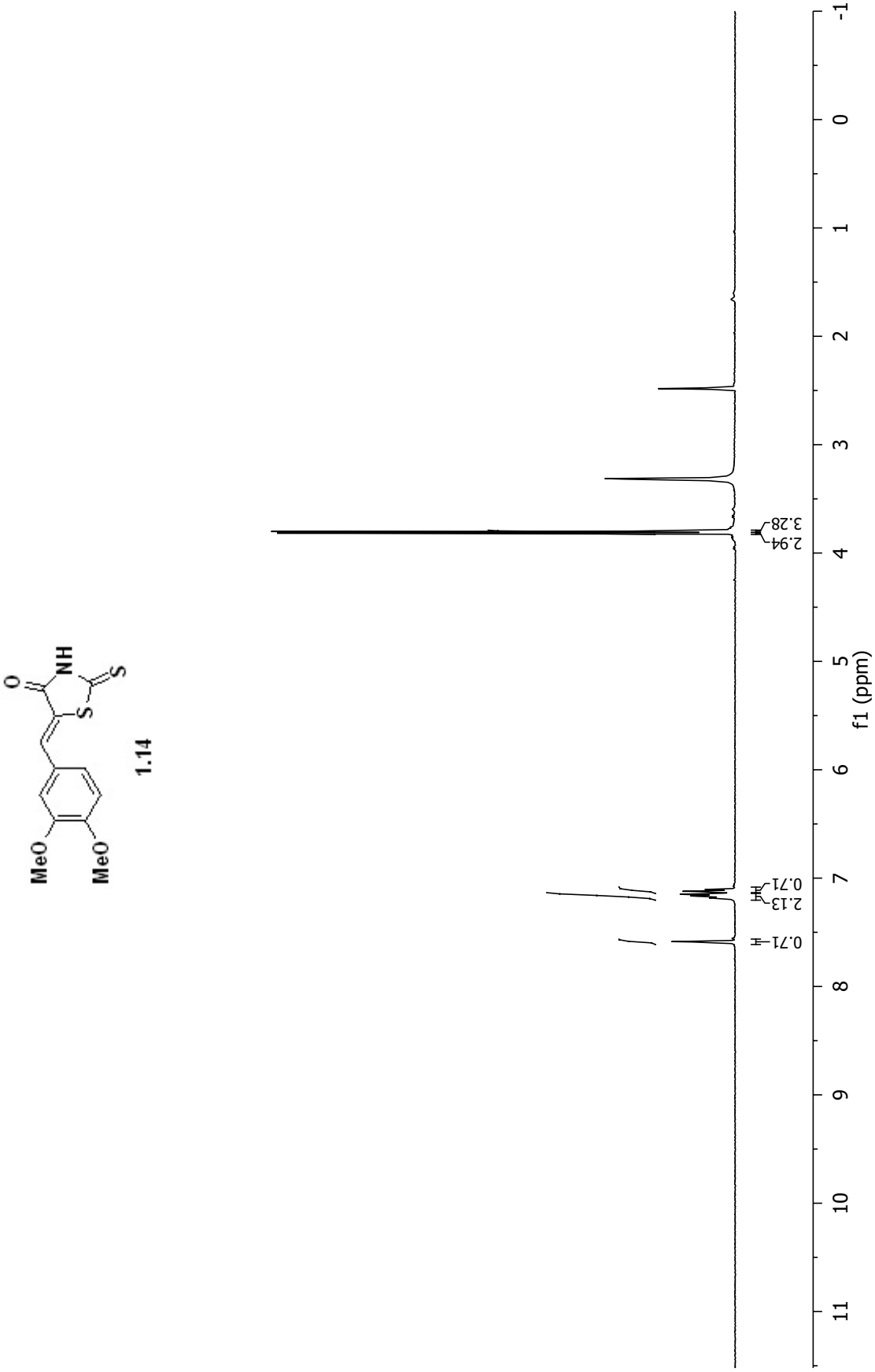

-10
0
10
20
30
40
50
60
70
80
90
100
110
120
130
140
150
160
170
180
190
200
f1 (ppm)

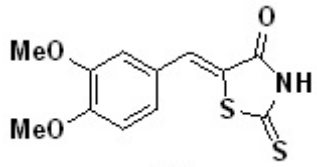

iCRT 3
f1 (ppm)


O
O
S
N
N
H

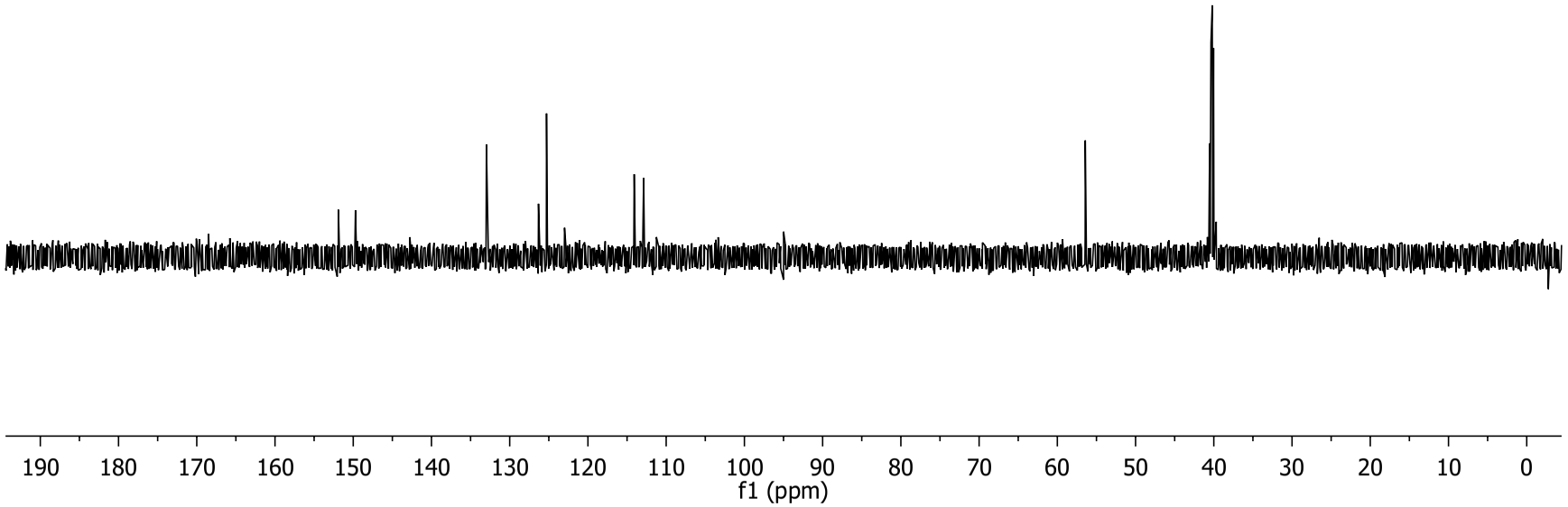
iCRT 3

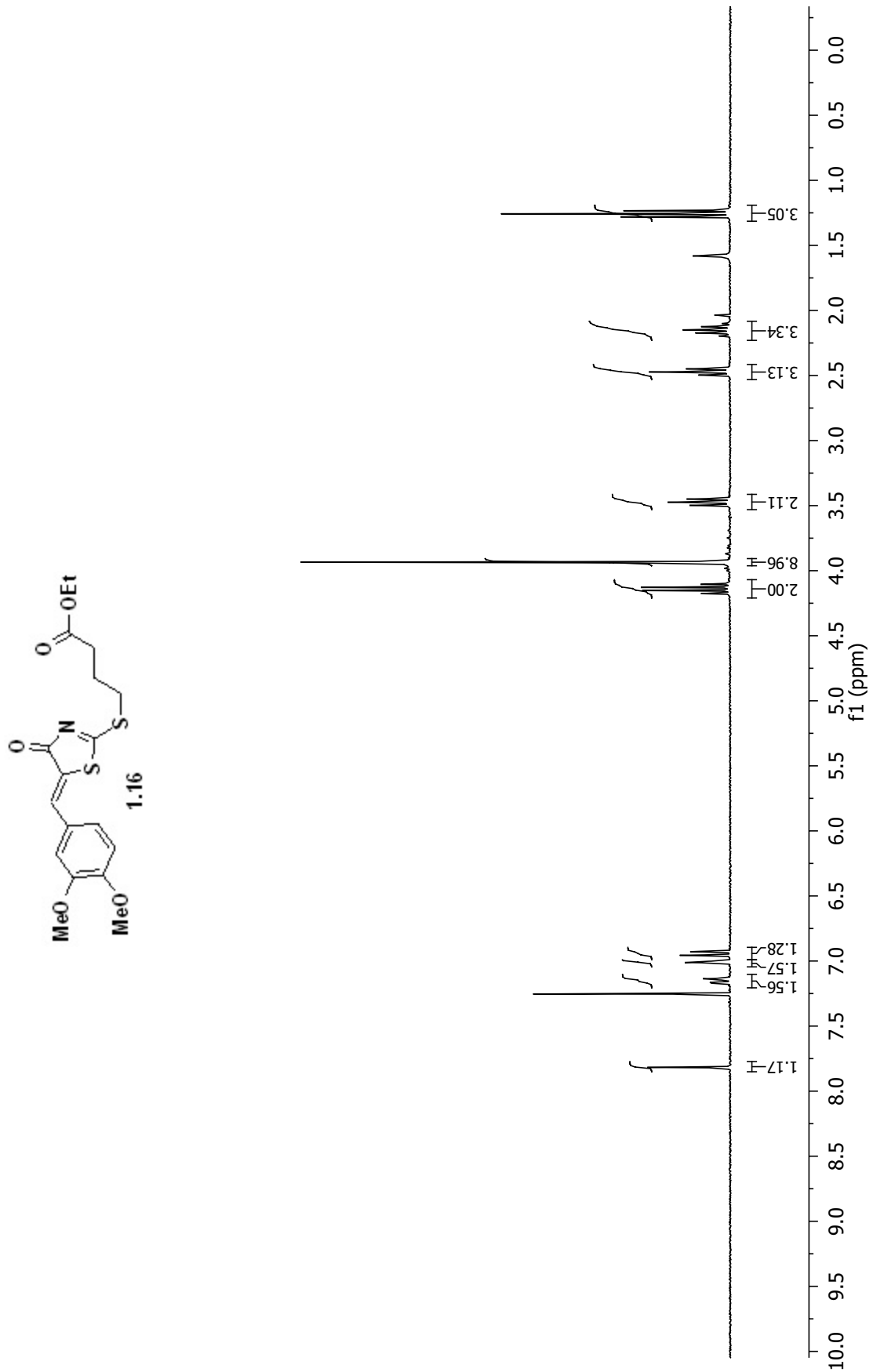

200
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
-10
f1 (ppm)

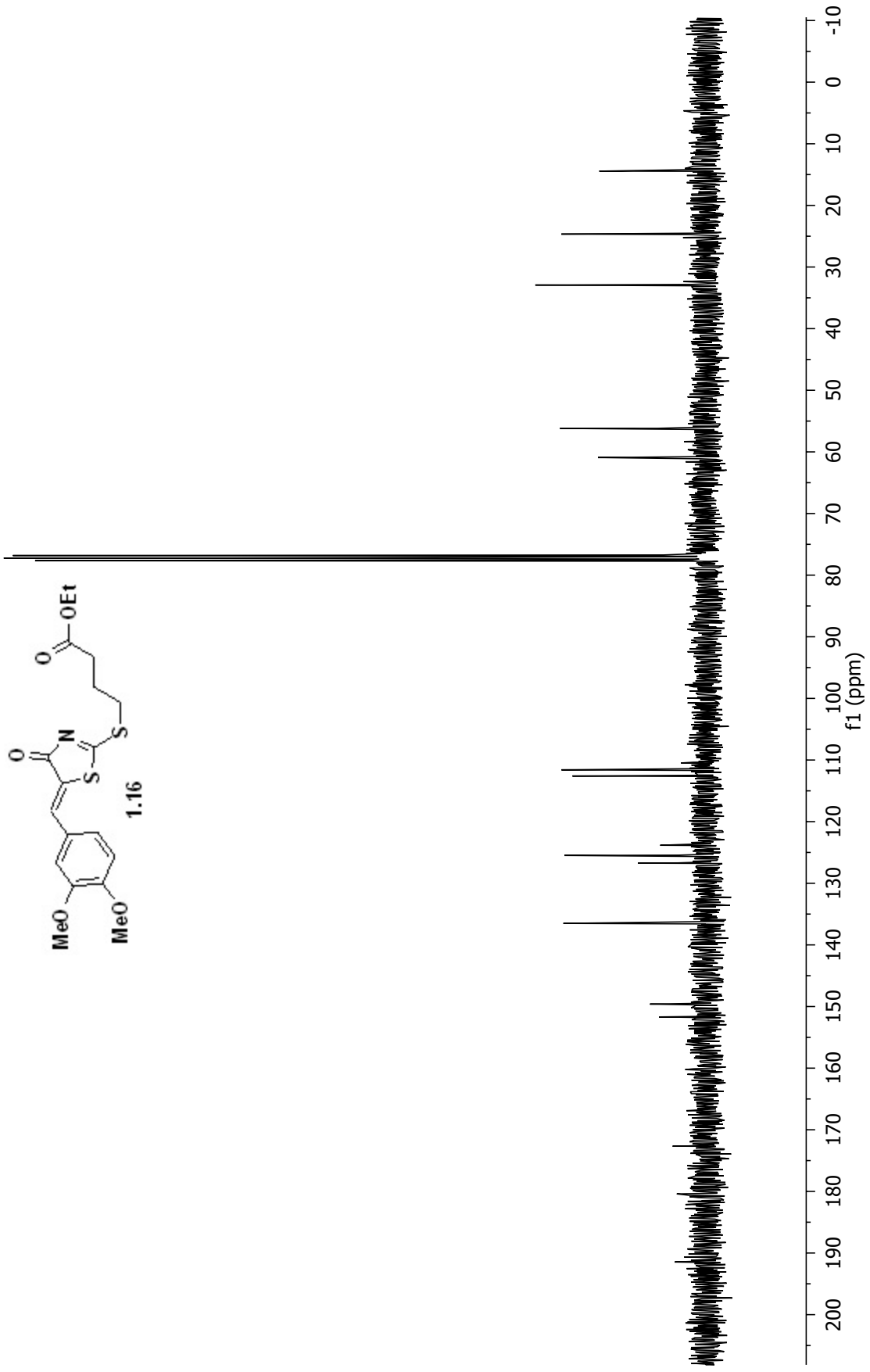

O
NH
S
S
MeO
MeO
1.14
2.94
3.28
2.13
0.71
0.71
f1 (ppm)

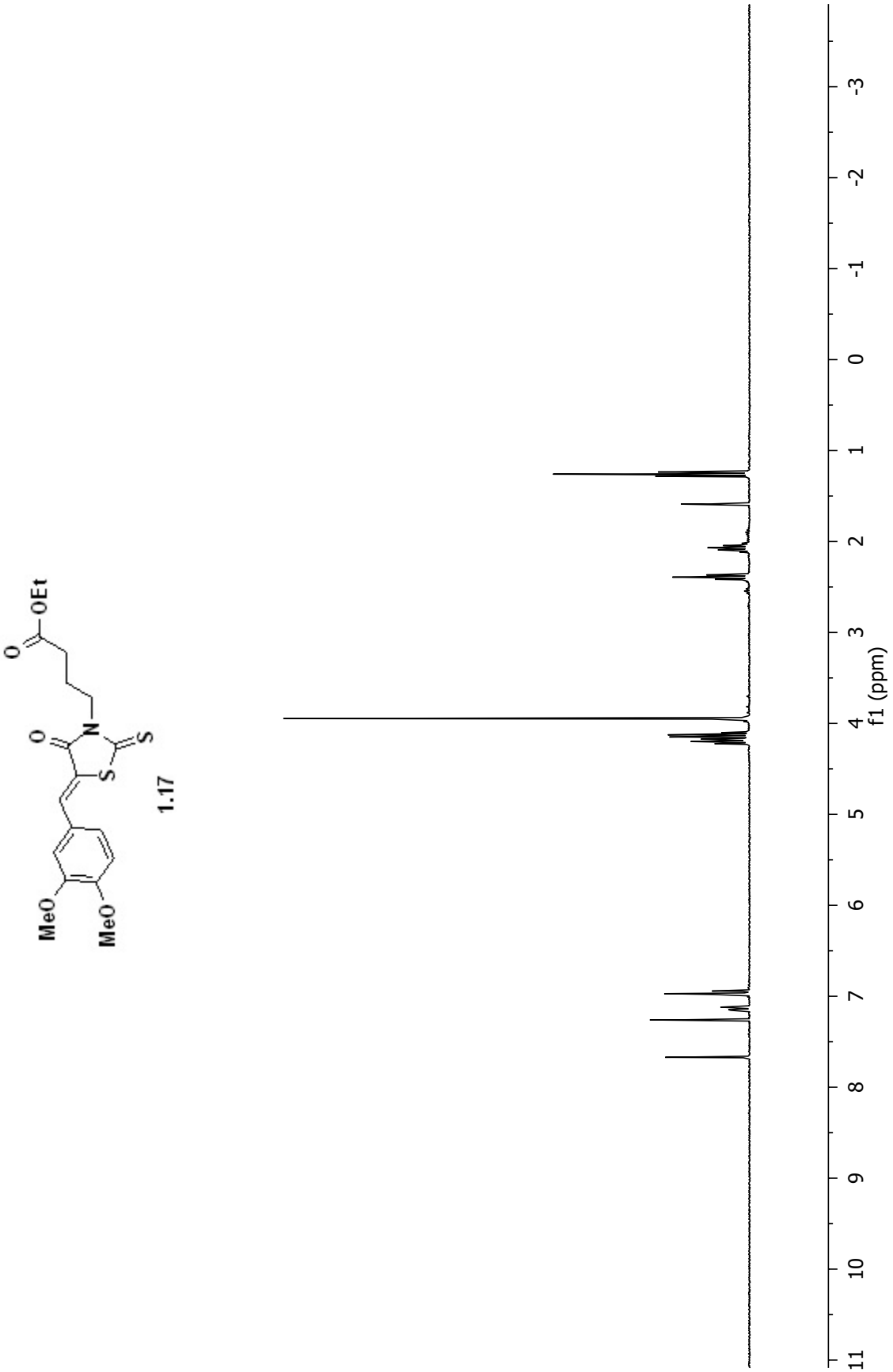

MeO
MeO
O
NH
S
S

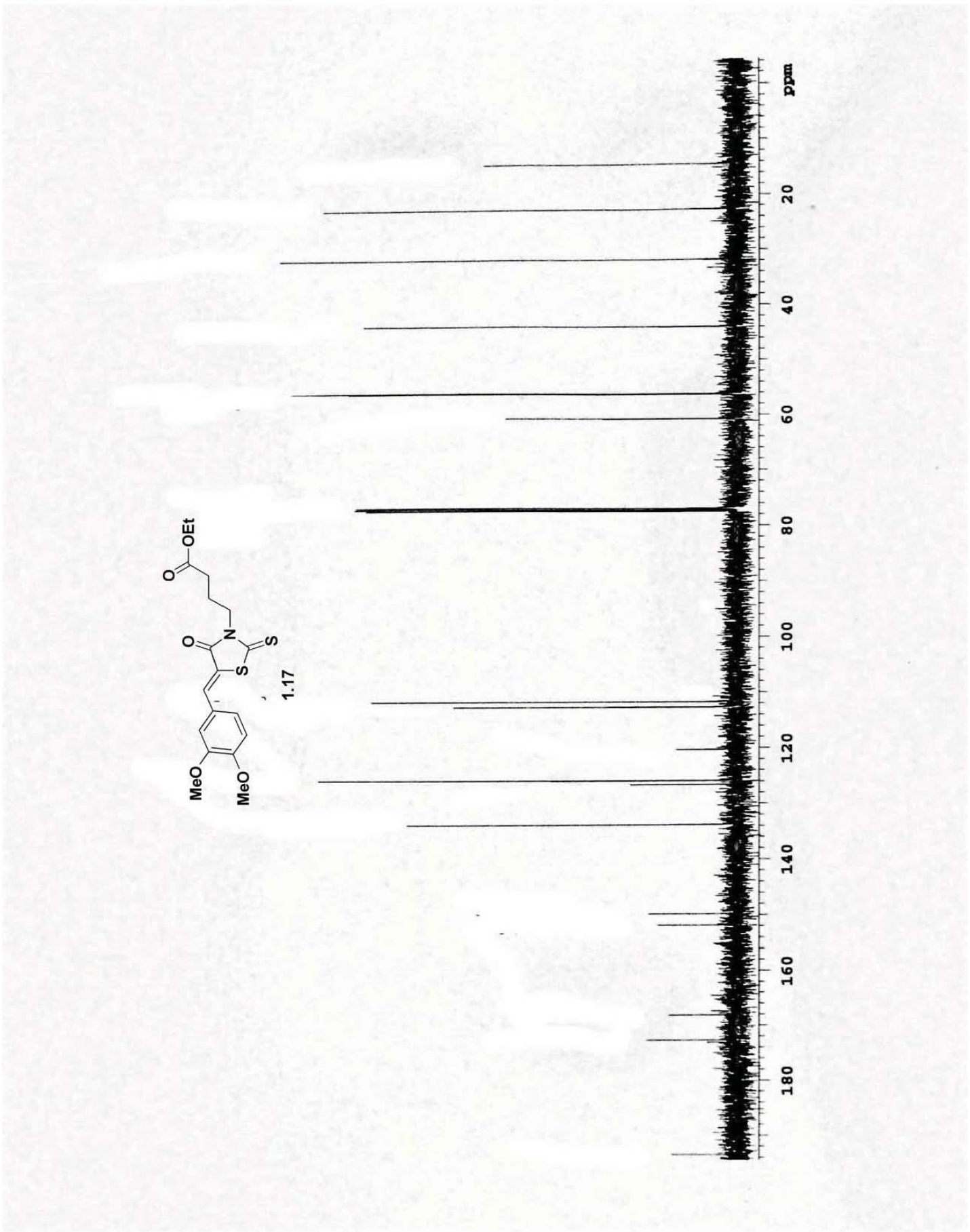
1.14

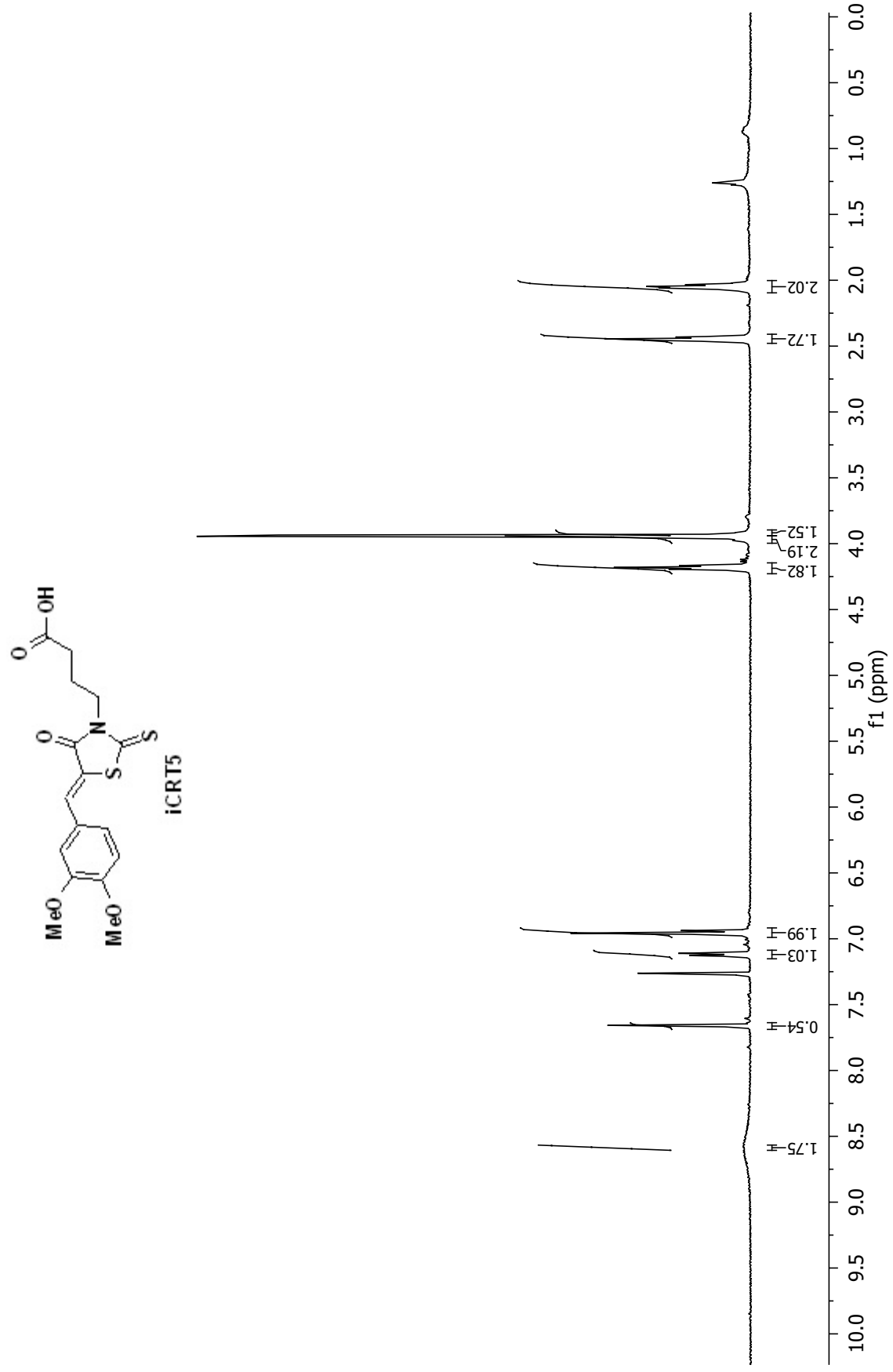

190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
f1 (ppm)

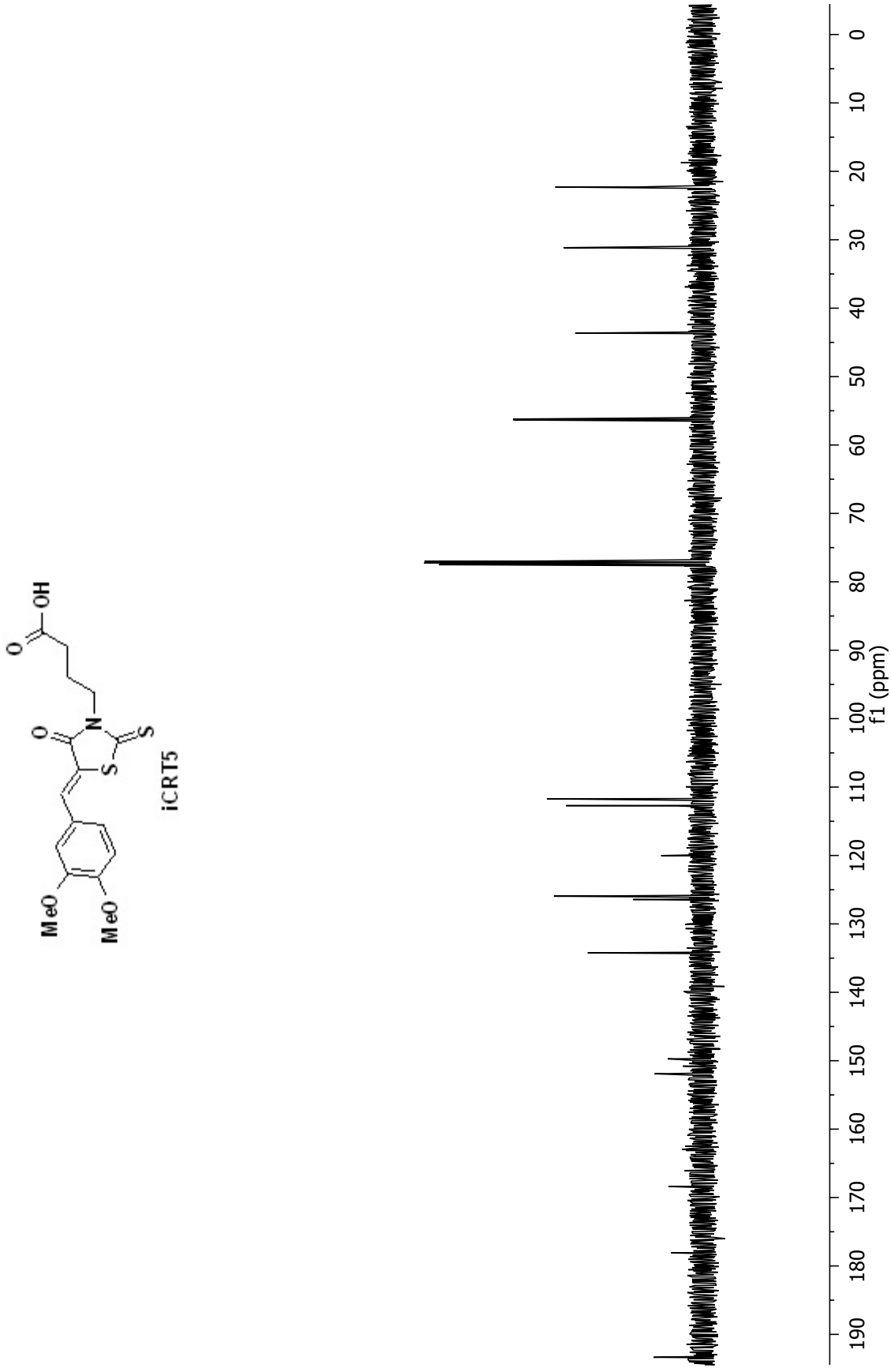

OEt
O
N
S
S
O
MeO
MeO
1.16
3.05
3.34
3.13
2.11
8.96
2.00
1.17
1.56
1.57
1.28
10.0
9.5
9.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
f1 (ppm)

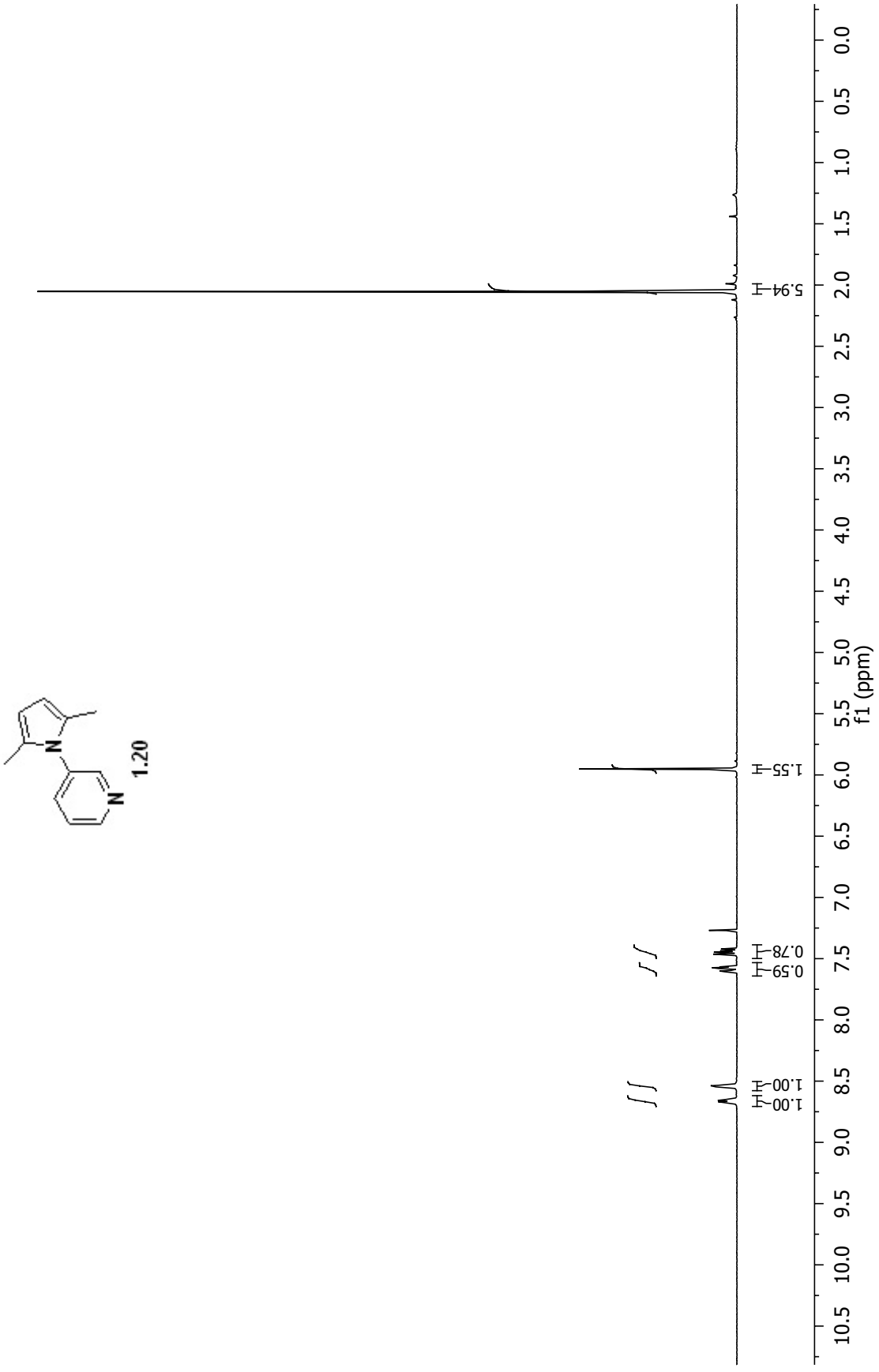

OEt
O
N
S
S
O
MeO
MeO
1.16
f1 (ppm)

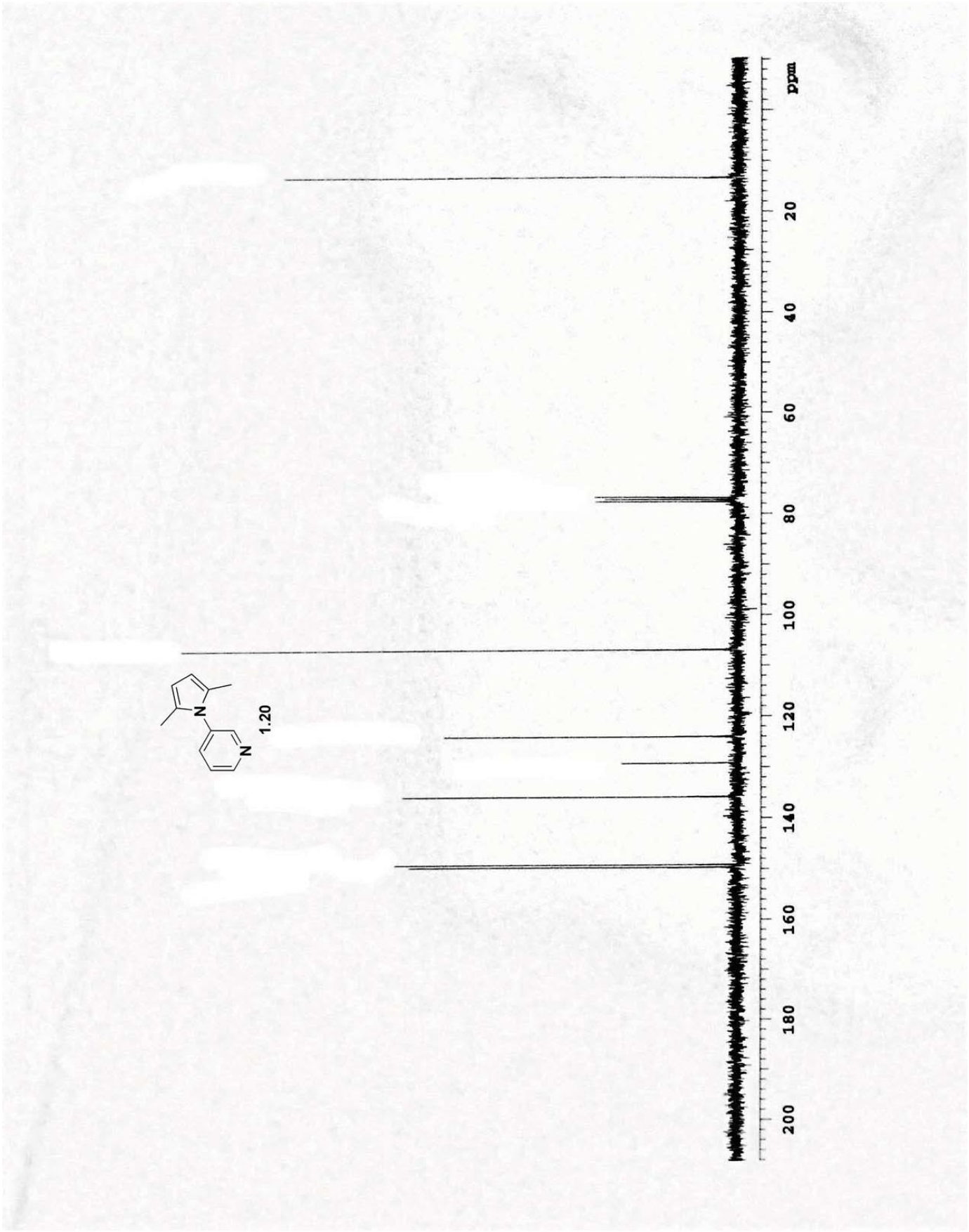

O
OEt
O
N
S
S
MeO
MeO
1.17
f1 (ppm)
11
10
9
8
7
6
5
4
3
2
1
0
-1
-2
-3

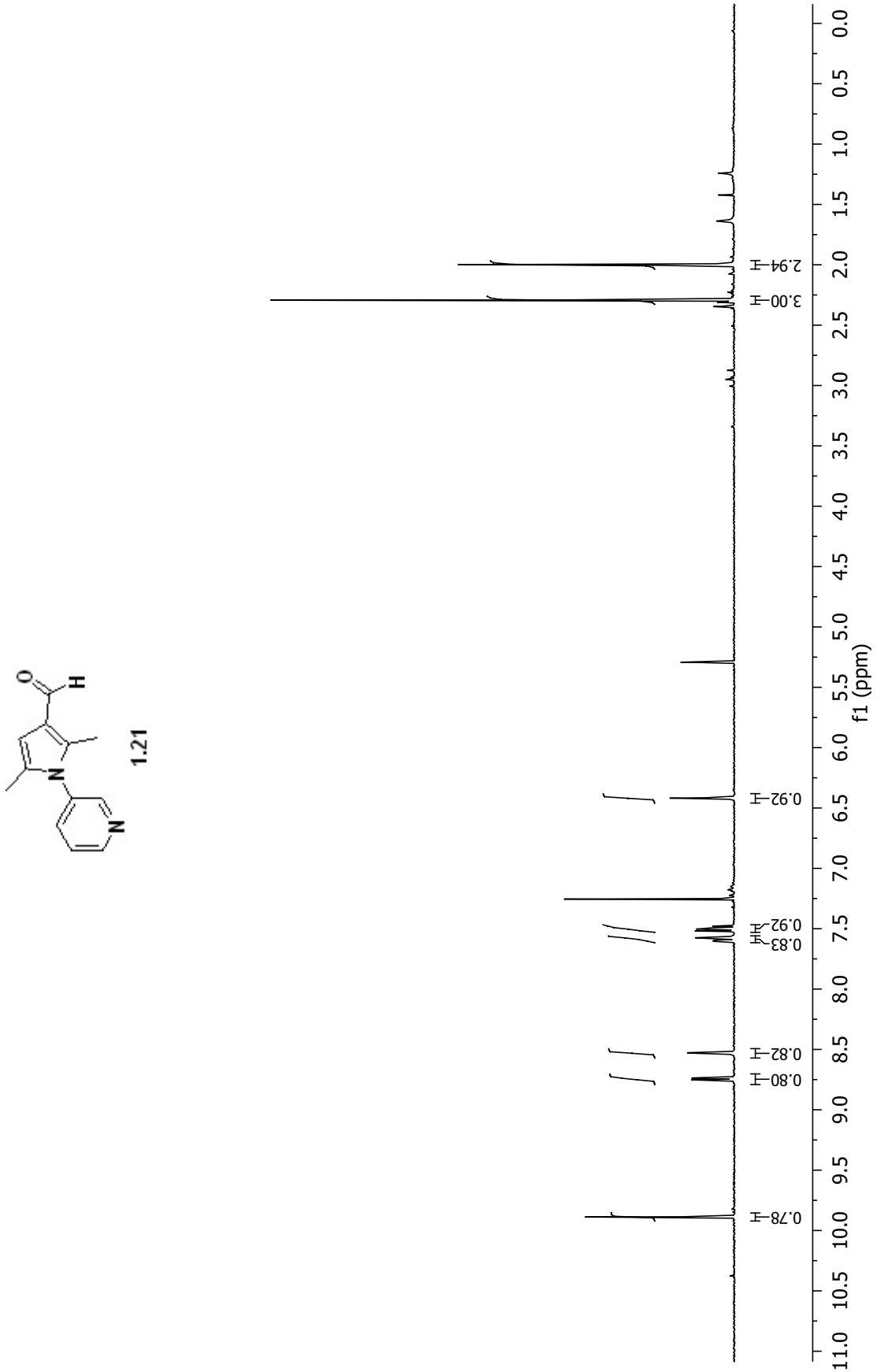

OEt
O
O
N
S
S
MeO
MeO
1.17
ppm
20
40
60
80
100
120
140
160
180

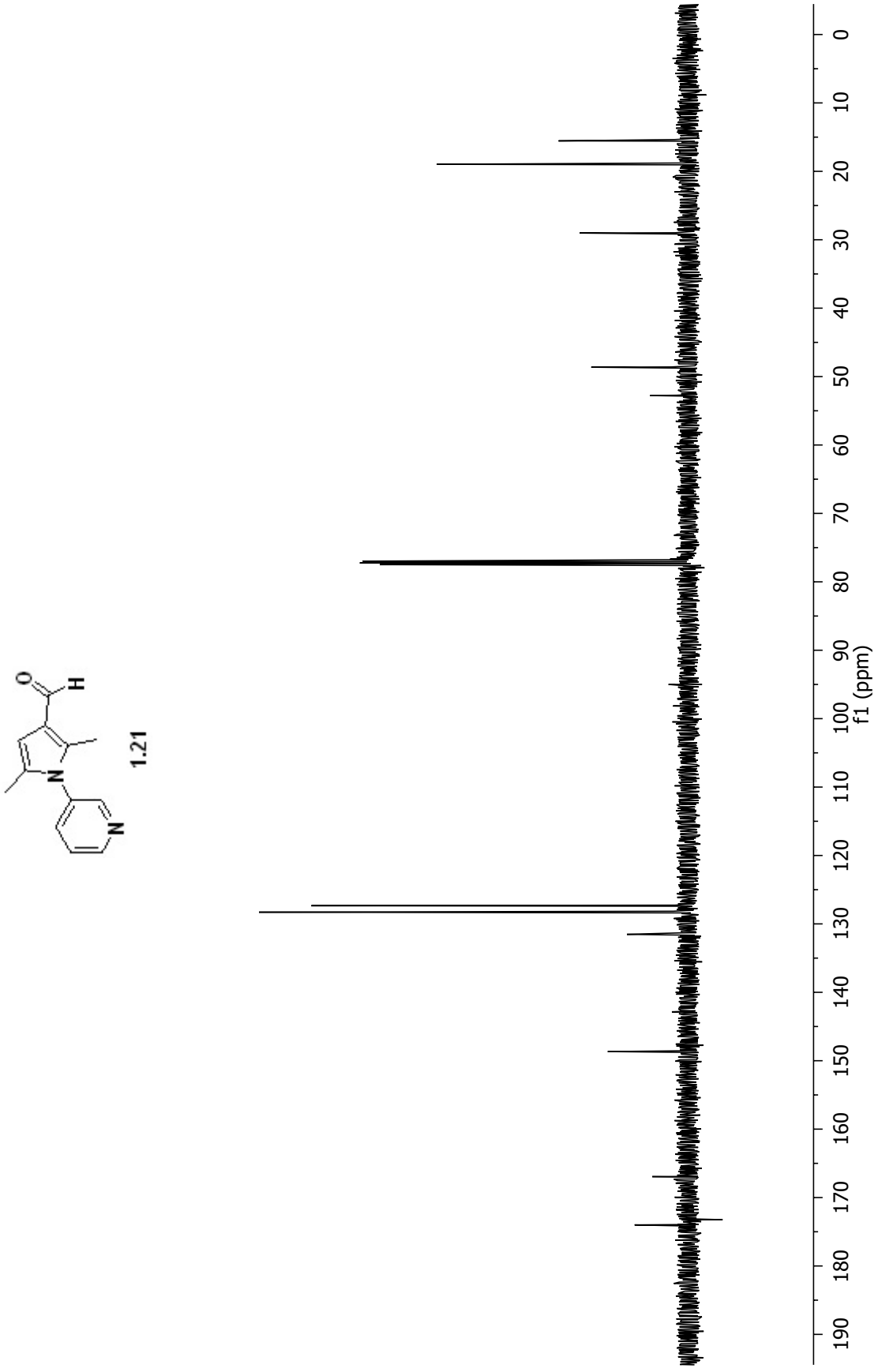

iCRT5
1.75
0.54
1.03
1.99
1.82
2.19
1.52
1.72
2.02
f1 (ppm)

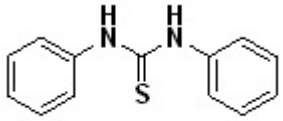

O
OH
O
N
S
S
MeO
MeO
iCRT5
f1 (ppm)
0
10
20
30
40
50
60
70
80
90
100
110
120
130
140
150
160
170
180
190


1.20
5.94
1.55
0.59
0.78
1.00
1.00
f1 (ppm)

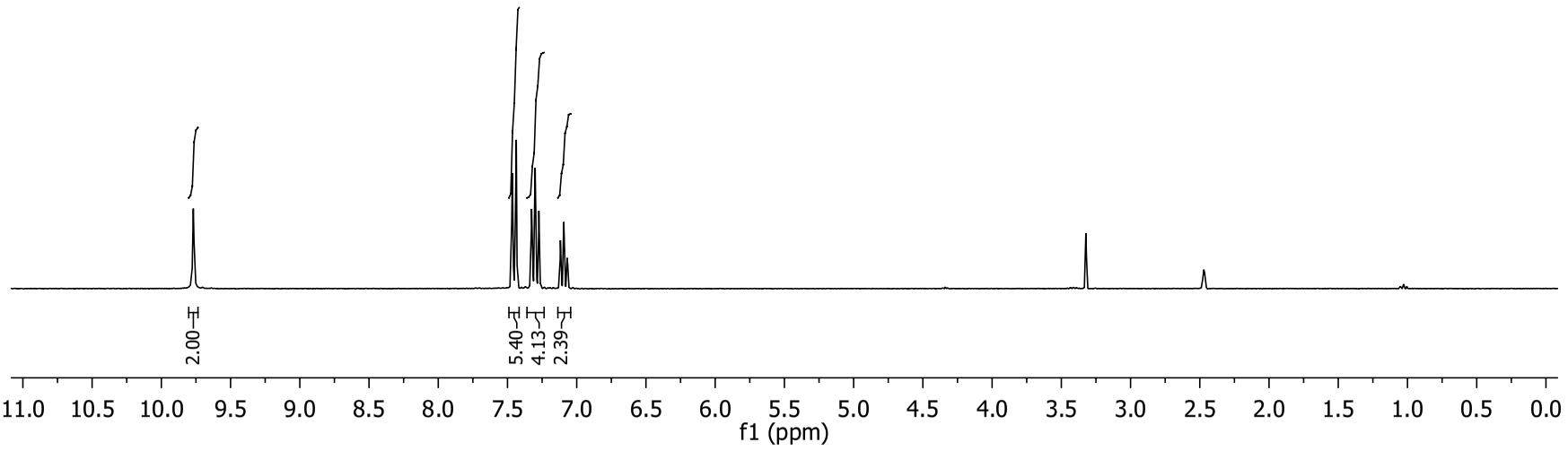

1.20
ppm
200
180
160
140
120
100
80
60
40
20

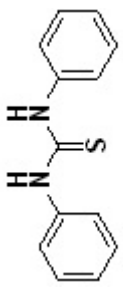

O
H
N
N
1.21
0.78
0.80
0.82
0.83
0.92
0.92
2.94
3.00
f1 (ppm)


O
H
N
N
1.21
0
10
20
30
40
50
60
70
80
90
100
110
120
130
140
150
160
170
180
190
f1 (ppm)

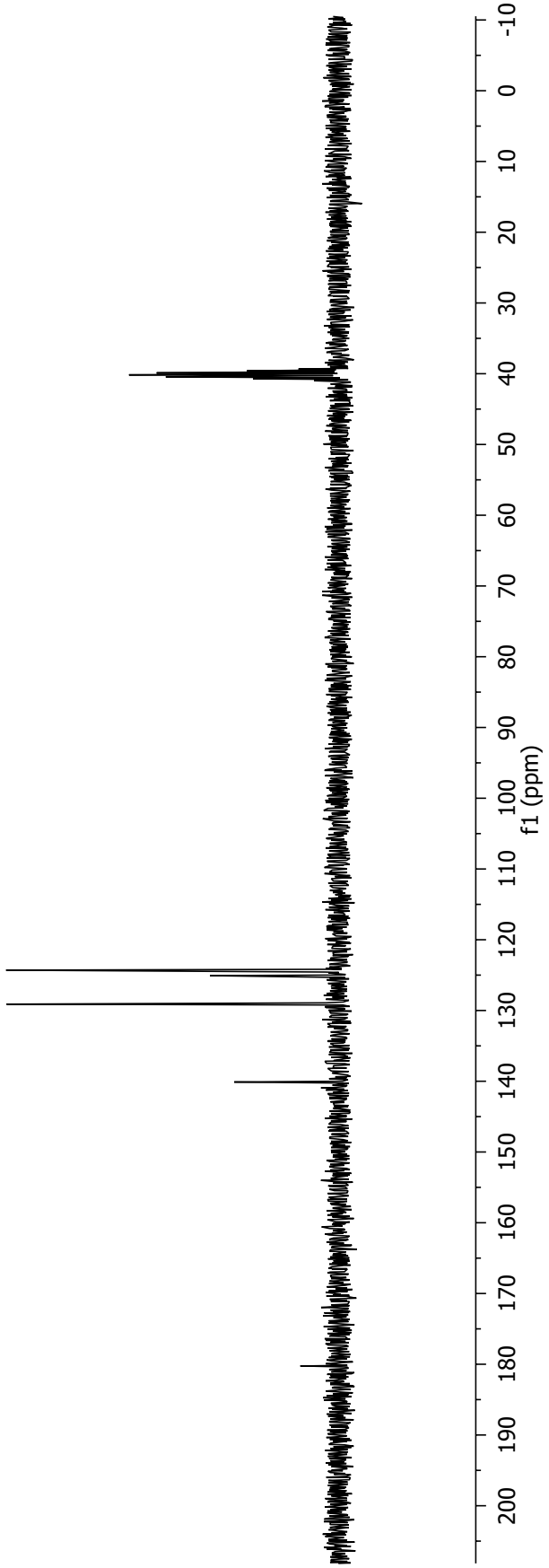


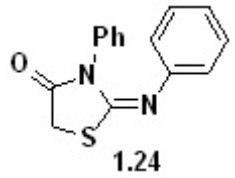
1.23


2.00
5.40
4.13
2.39
11.0 10.5 10.0 9.5 9.0 8.5 8.0 7.5 7.0 6.5 6.0 5.5 5.0 4.5 4.0 3.5 3.0 2.5 2.0 1.5 1.0 0.5 0.0
f1 (ppm)

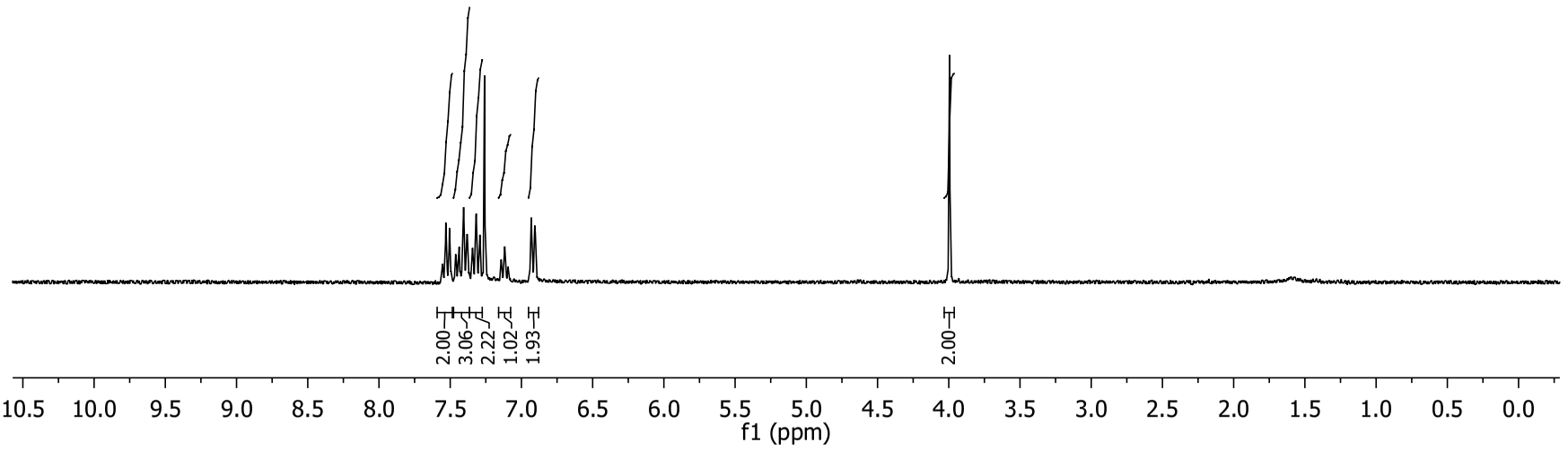


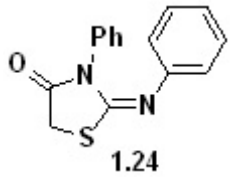
1.23


-10
0
10
20
30
40
50
60
70
80
90
100
110
120
130
140
150
160
170
180
190
200
f1 (ppm)

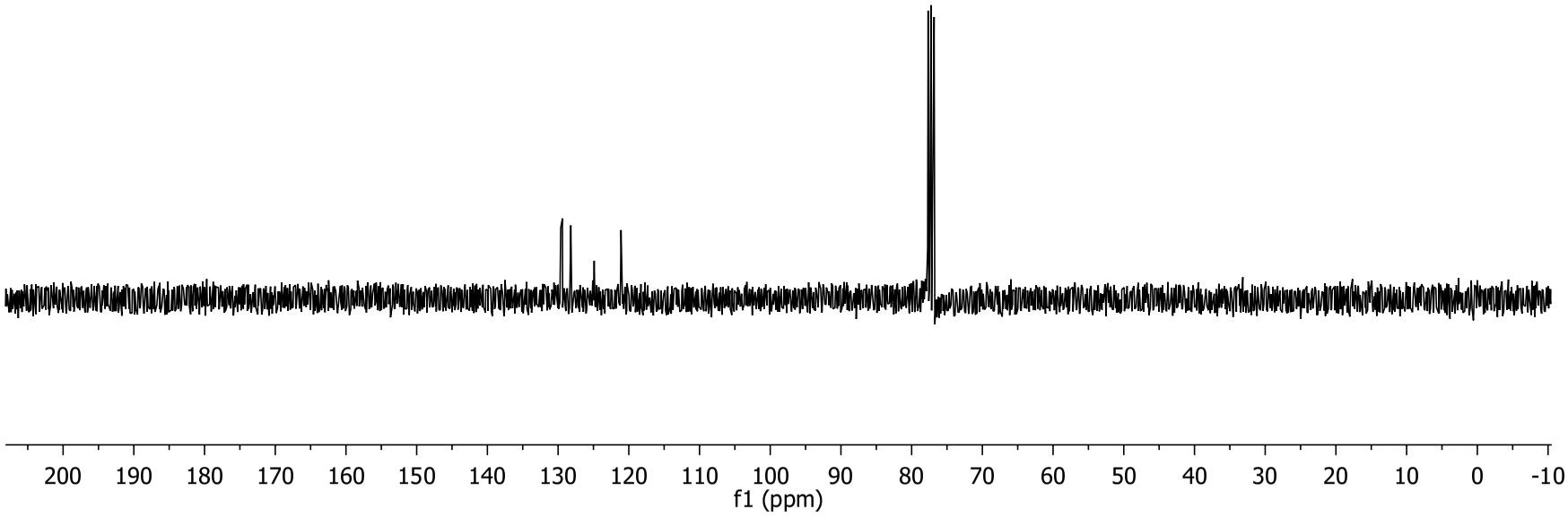

Ph
O
N
N
S

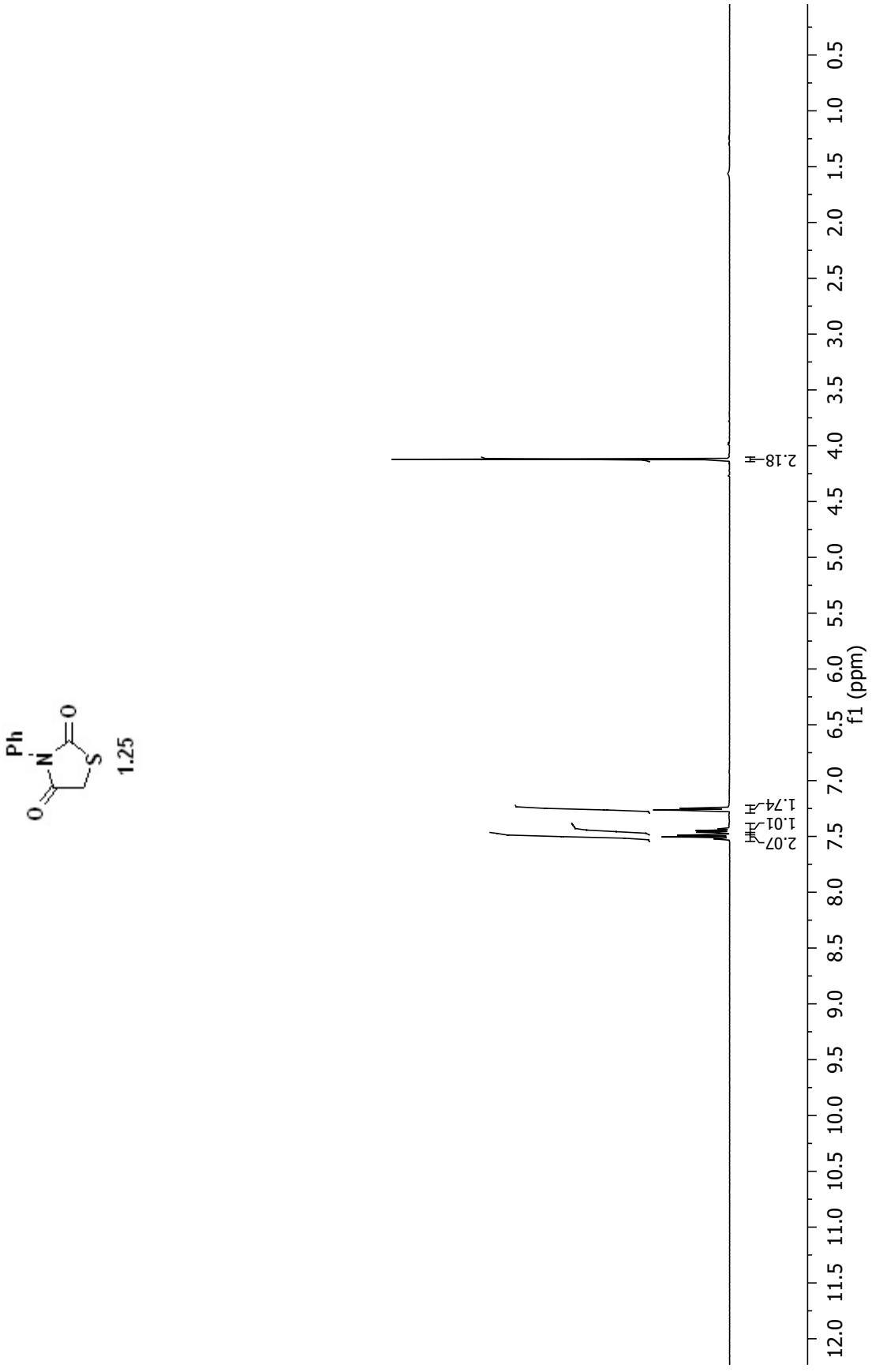
1.24

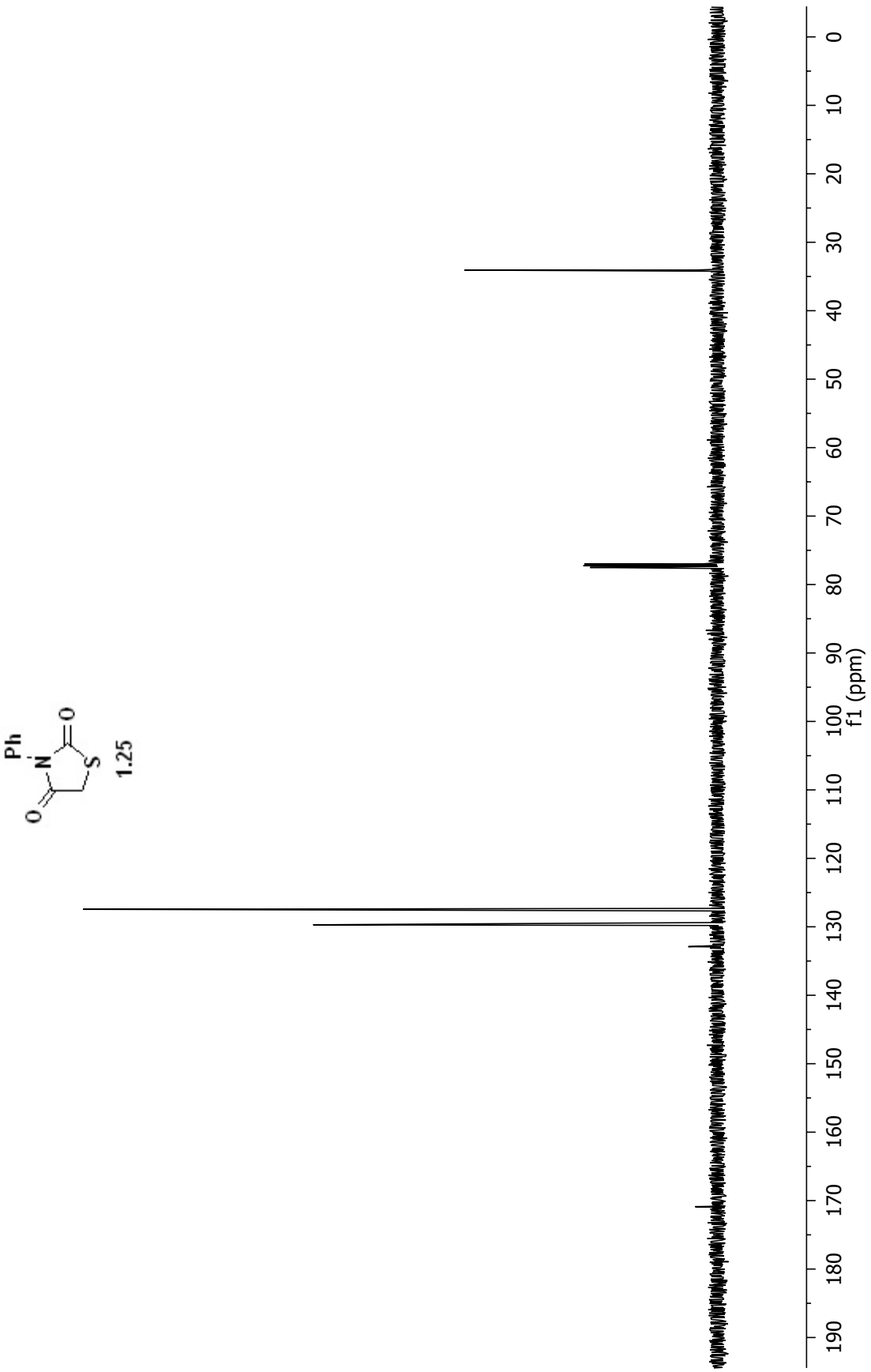

2.00
3.06
2.22
1.02
1.93
2.00
10.5 10.0 9.5 9.0 8.5 8.0 7.5 7.0 6.5 6.0 5.5 5.0 4.5 4.0 3.5 3.0 2.5 2.0 1.5 1.0 0.5 0.0
f1 (ppm)

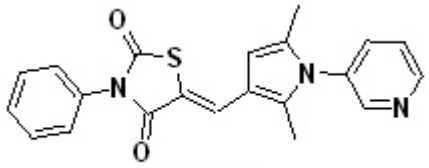

Ph
O
N
N
S

1.24

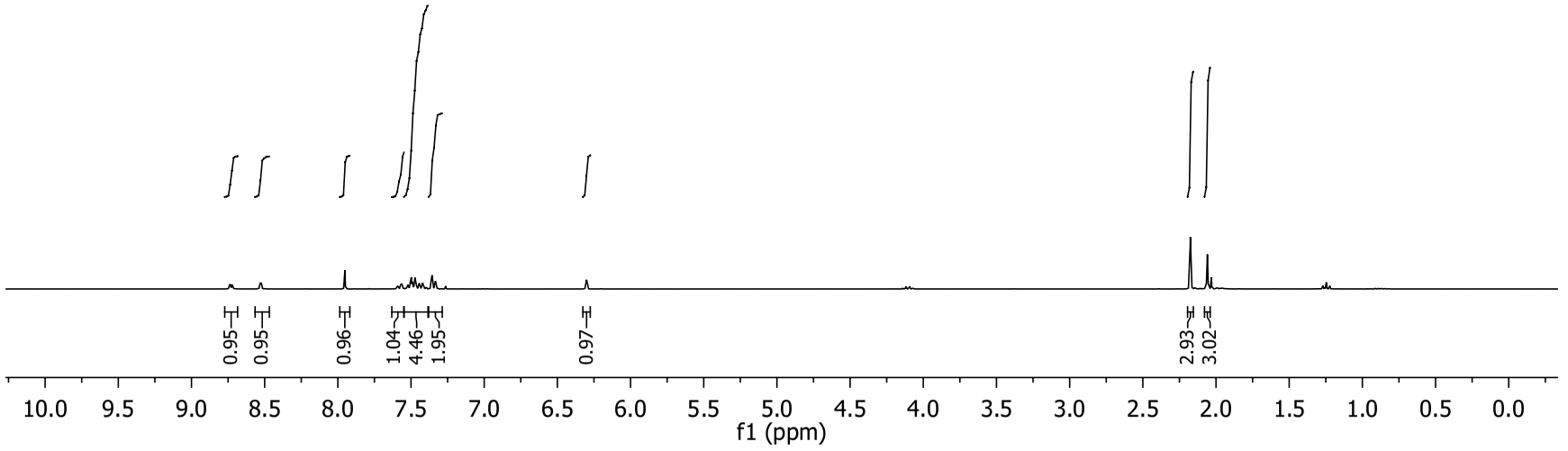

200
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
-10
f1 (ppm)

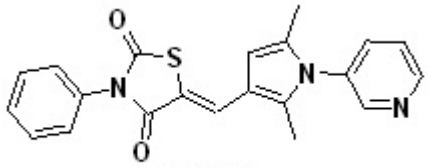

Ph
N
O
O
S
1.25
2.18
1.74
1.01
2.07
f1 (ppm)


Ph
N
O
O
S
O
1.25
f1 (ppm)

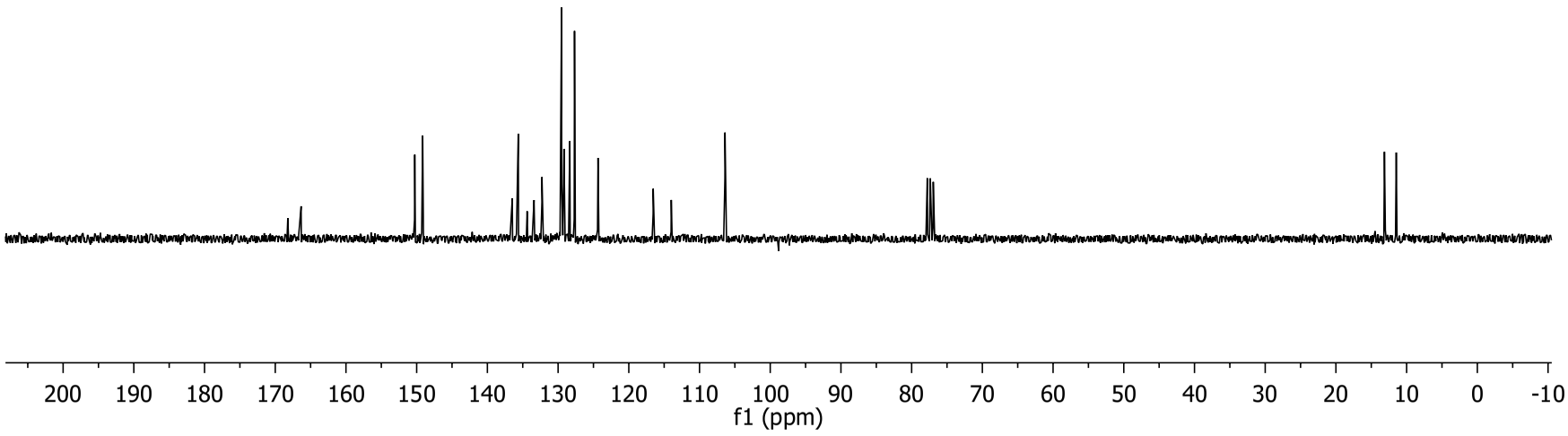

O
S
N
N
N
O

iCRT14


0.95
0.95
0.96
1.04
4.46
1.95
0.97
2.93
3.02
10.0 9.5 9.0 8.5 8.0 7.5 7.0 6.5 6.0 5.5 5.0 4.5 4.0 3.5 3.0 2.5 2.0 1.5 1.0 0.5 0.0
f1 (ppm)

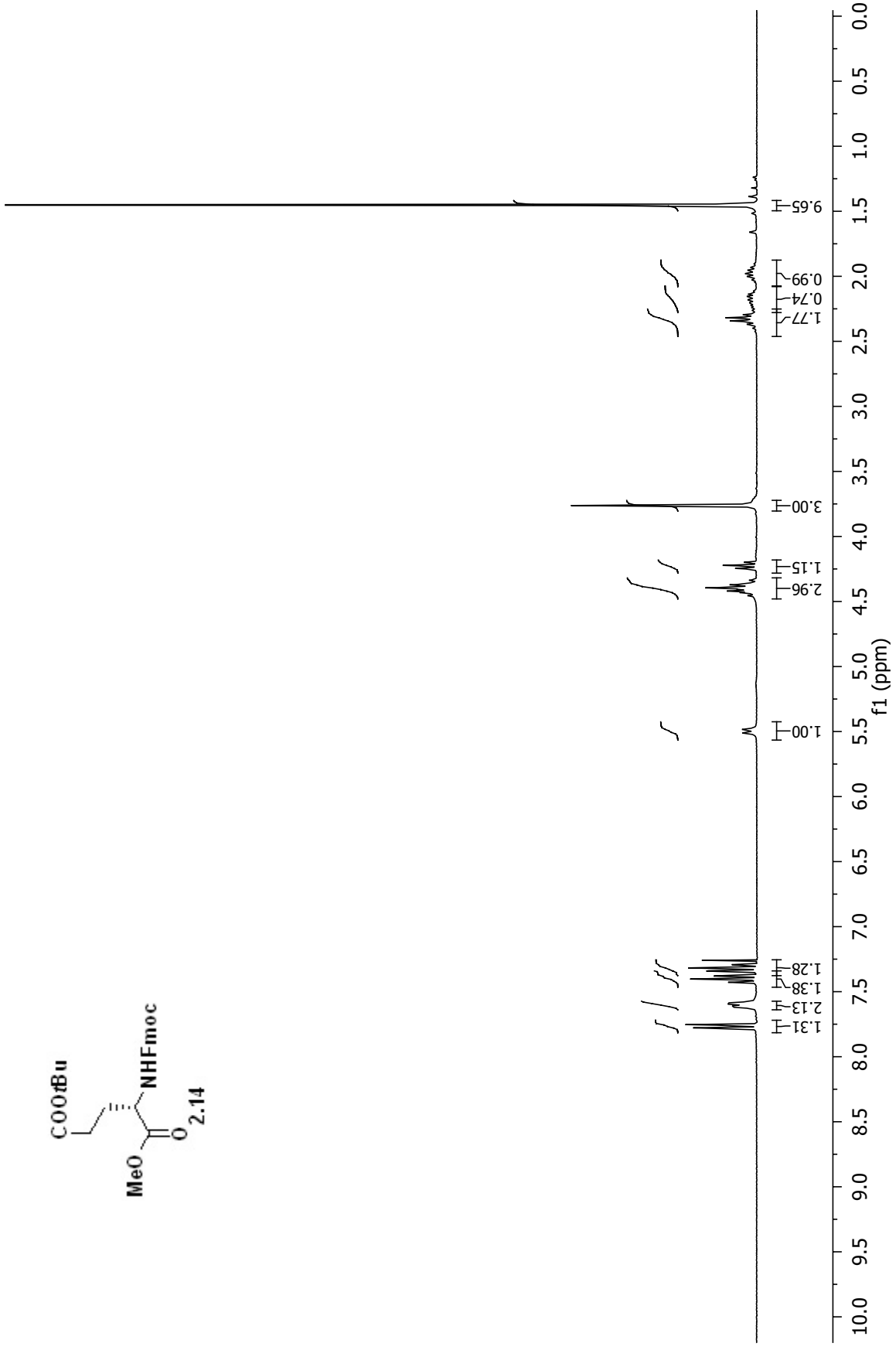

O
S
N
N
N
O

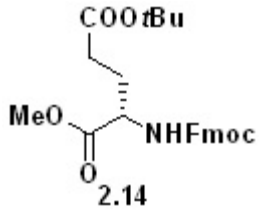
iCRT14


200
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
-10
f1 (ppm)

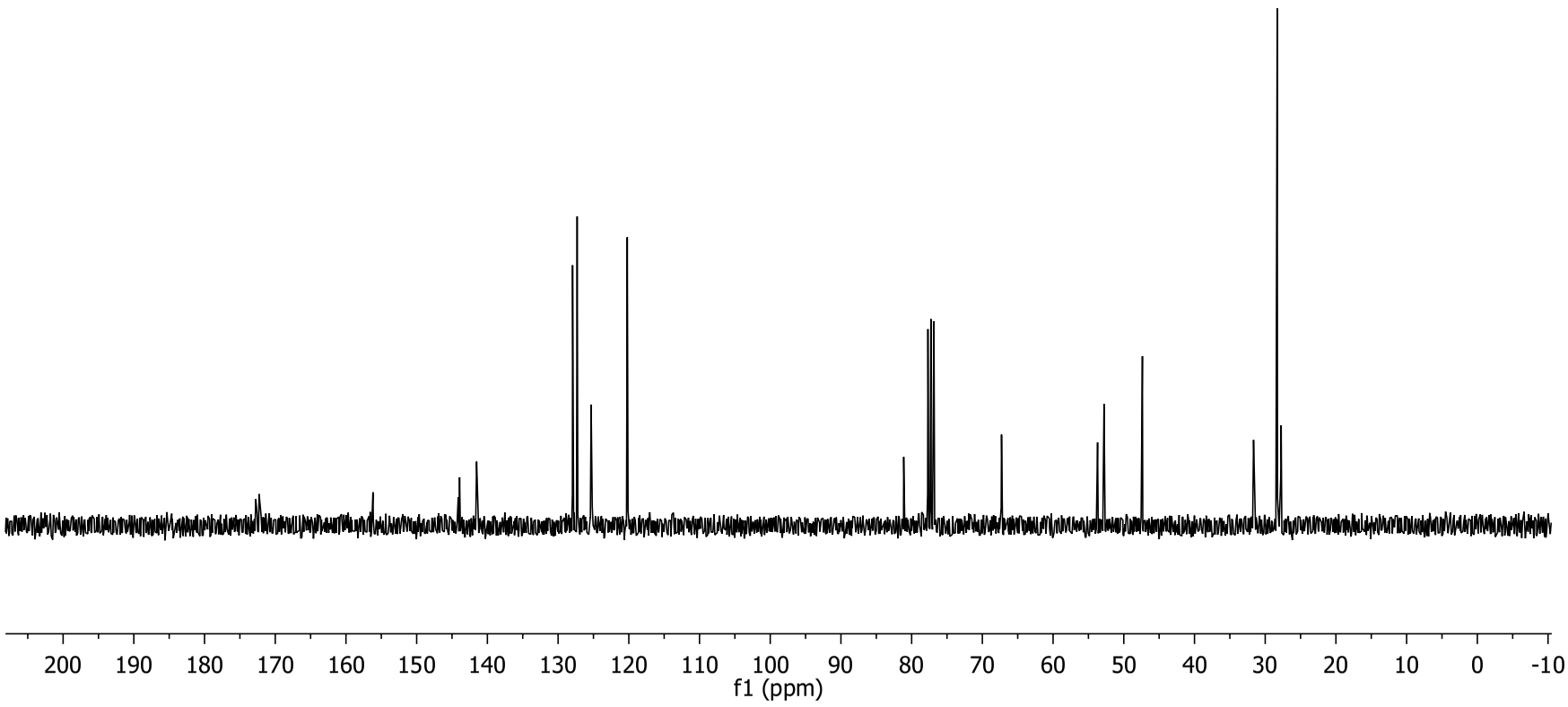
APPENDIX B

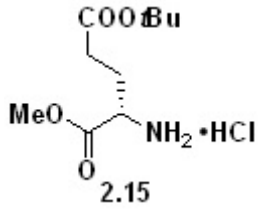
^{1}H AND ^{13}C NMR SPECTRA OF CHAPTER 2


COOtBu
NHFmoc
MeO
O
2.14
f1 (ppm)
1.31
2.13
1.38
1.28
1.00
2.96
1.15
3.00
1.77
0.74
0.99
9.65

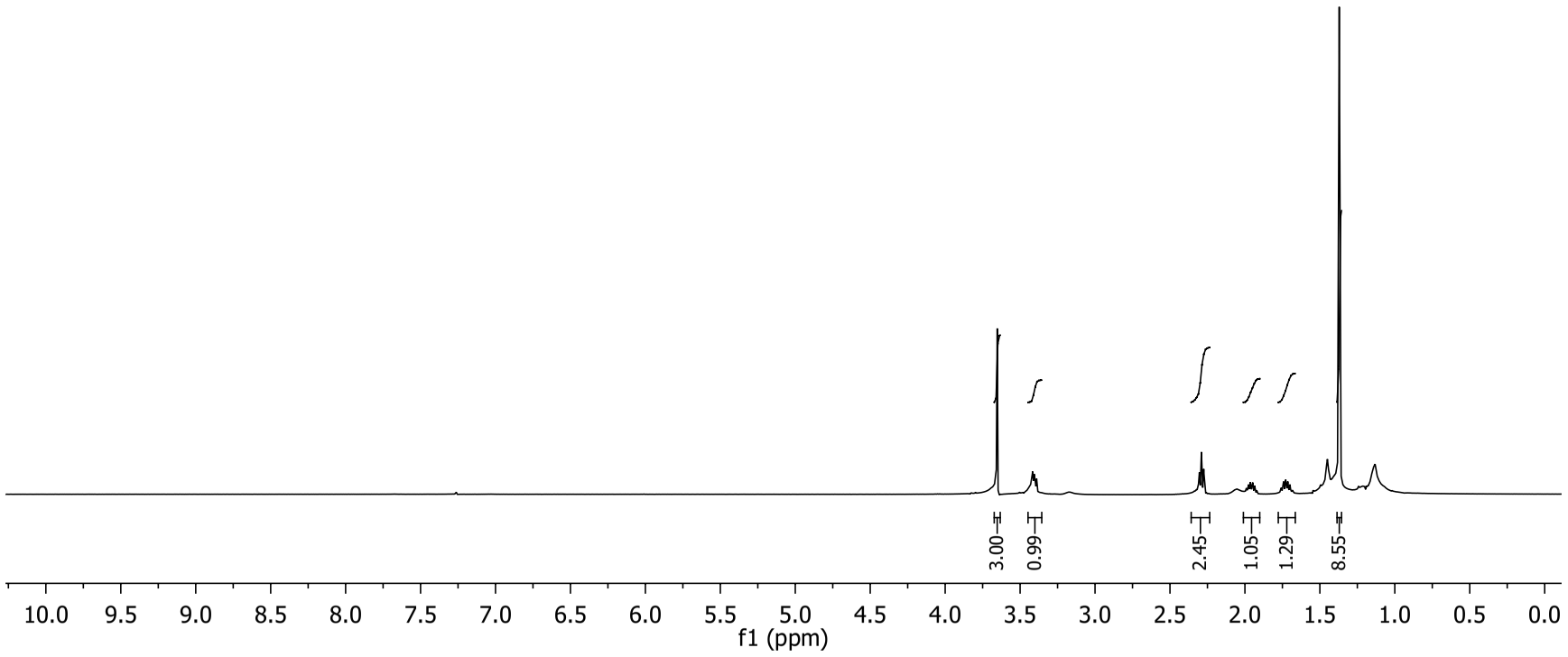

COOtBu
MeO
NHFmoc
O

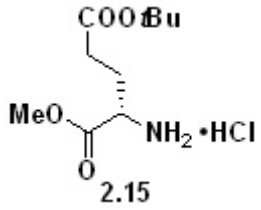
2.14


200 190 180 170 160 150 140 130 120 110 100 90 80 70 60 50 40 30 20 10 0 -10
f1 (ppm)

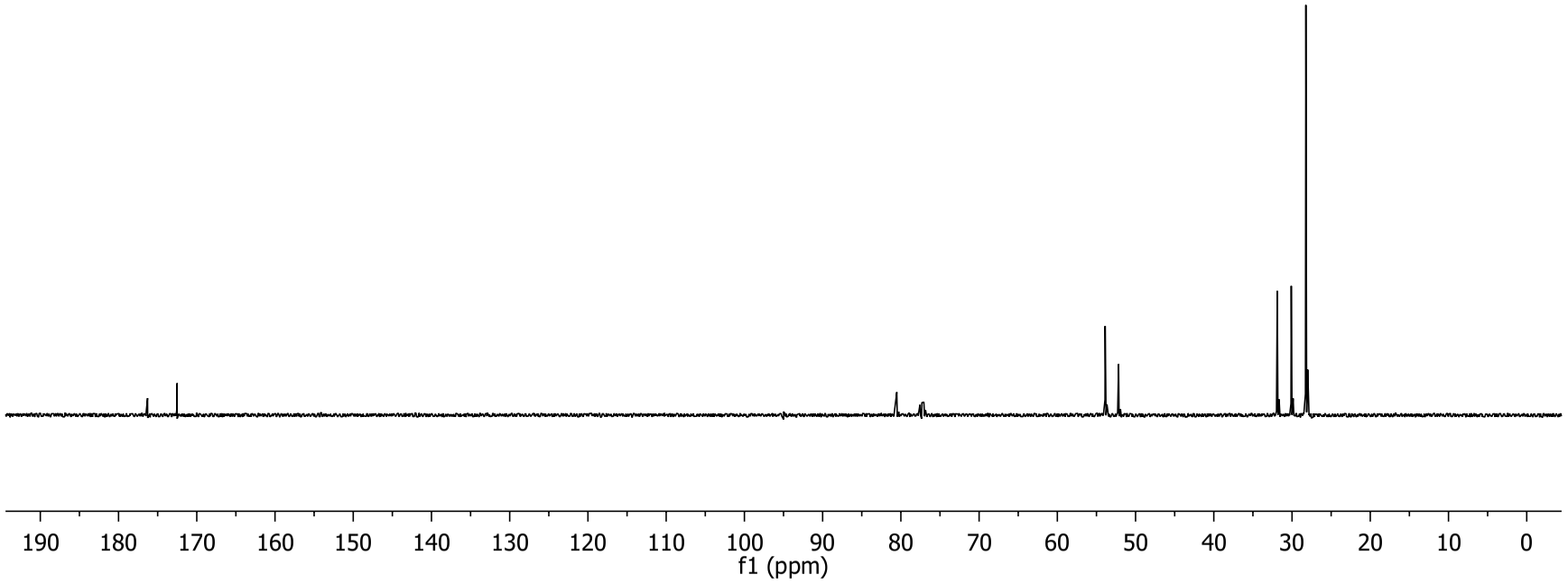

COOtBu
MeO
NH2•HCl
O

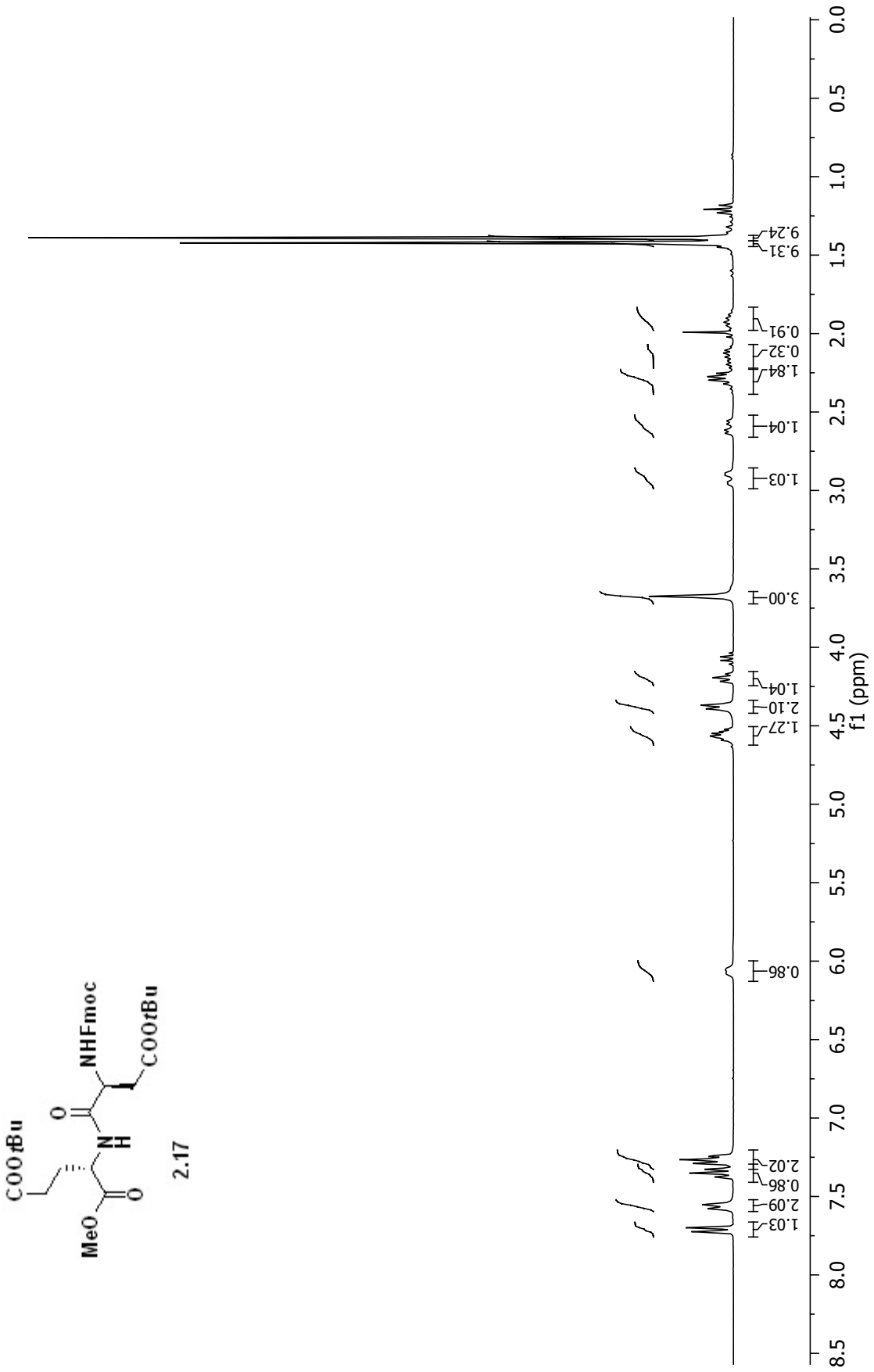
2.15

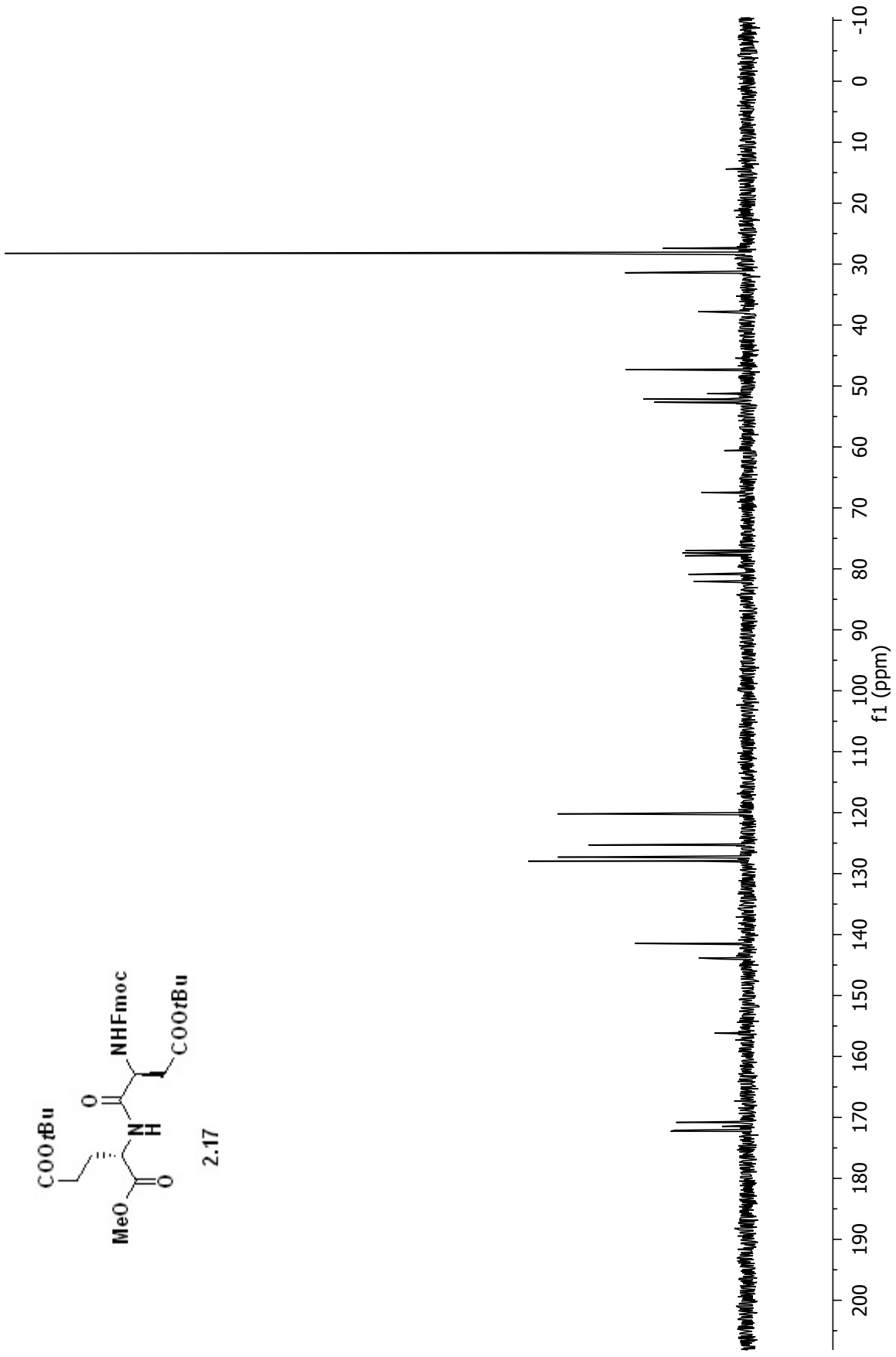

3.00
0.99
2.45
1.05
1.29
8.55
10.0 9.5 9.0 8.5 8.0 7.5 7.0 6.5 6.0 5.5 5.0 4.5 4.0 3.5 3.0 2.5 2.0 1.5 1.0 0.5 0.0
f1 (ppm)

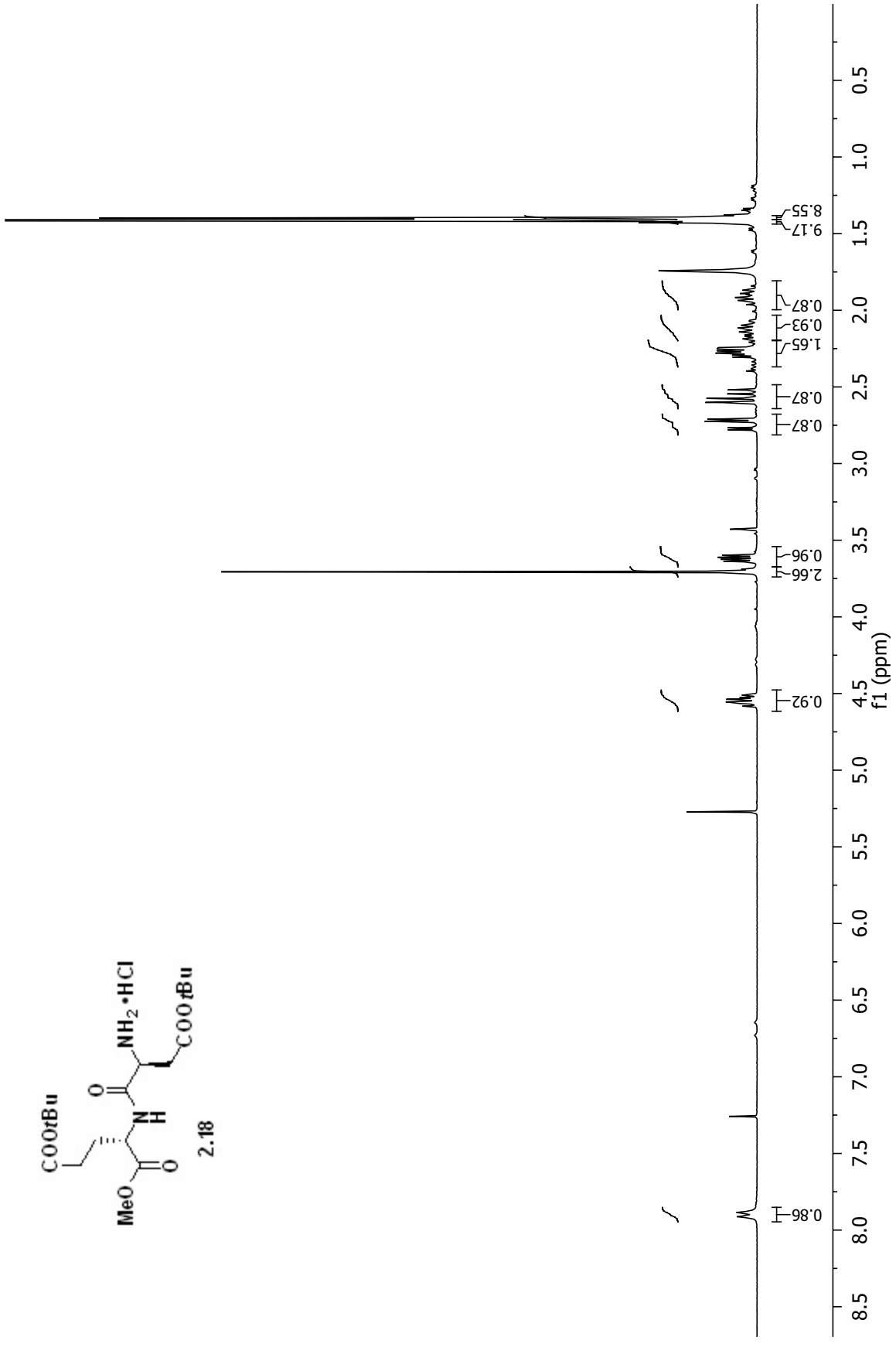

COOtBu
MeO
NH2•HCl
O

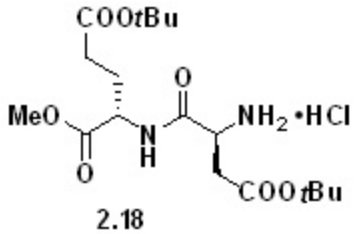
2.15


190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
f1 (ppm)

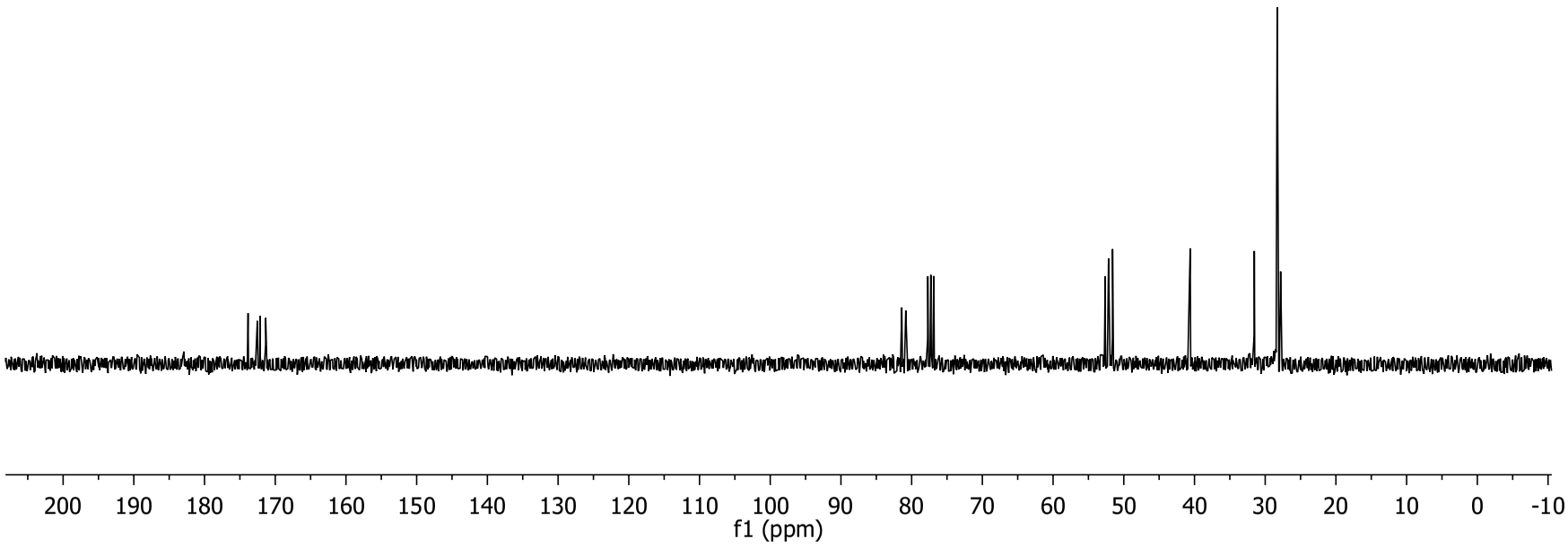

COOtBu
MeO
O
O
N
H
O
NHFmoc
COOtBu
2.17
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
f1 (ppm)
1.03
2.09
0.86
2.02
0.86
1.27
2.10
1.04
3.00
1.03
1.04
1.84
0.32
0.91
9.31
9.24

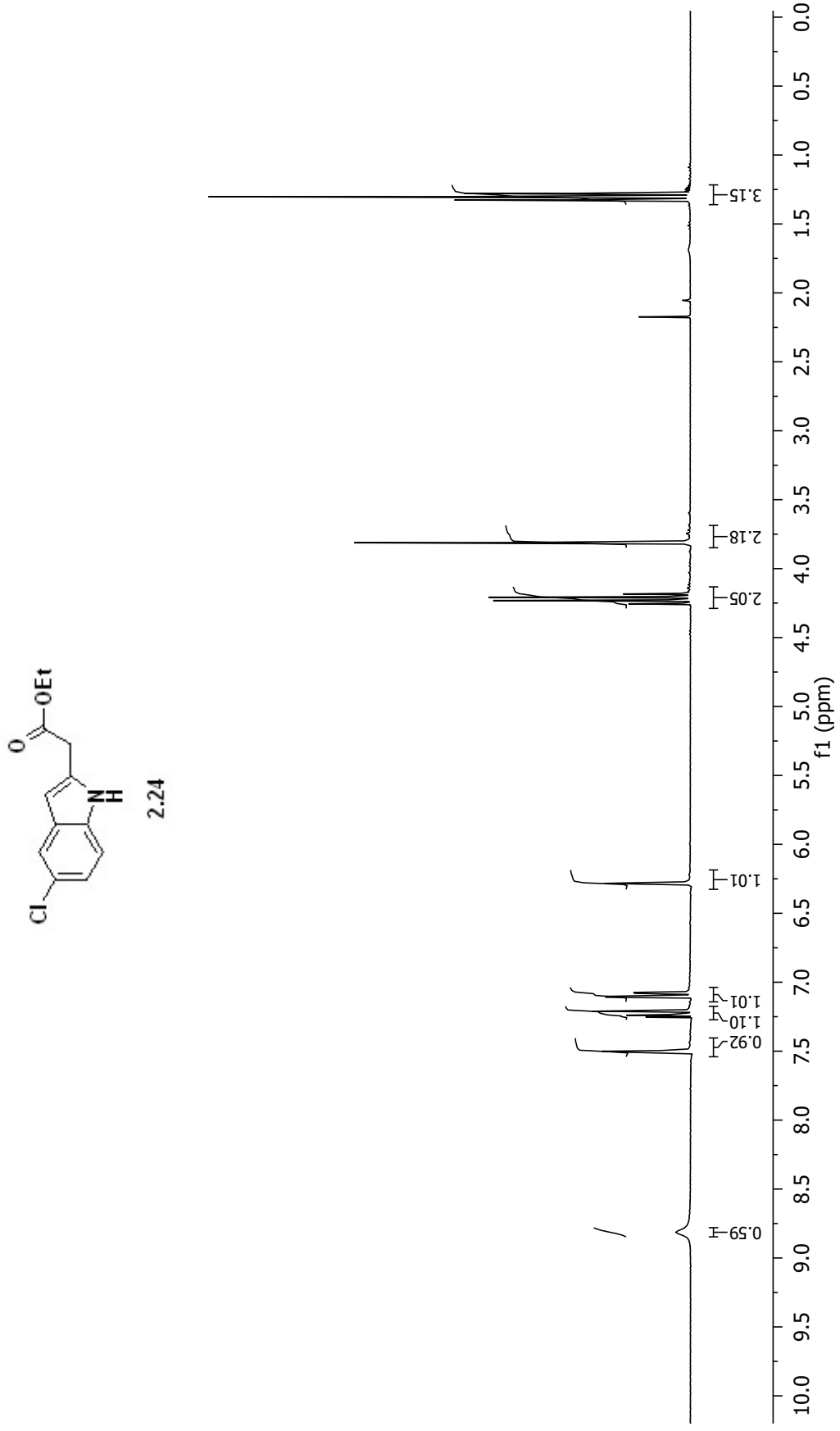

COOtBu
MeO
O
N
H
O
NHFmoc
COOtBu
2.17
f1 (ppm)

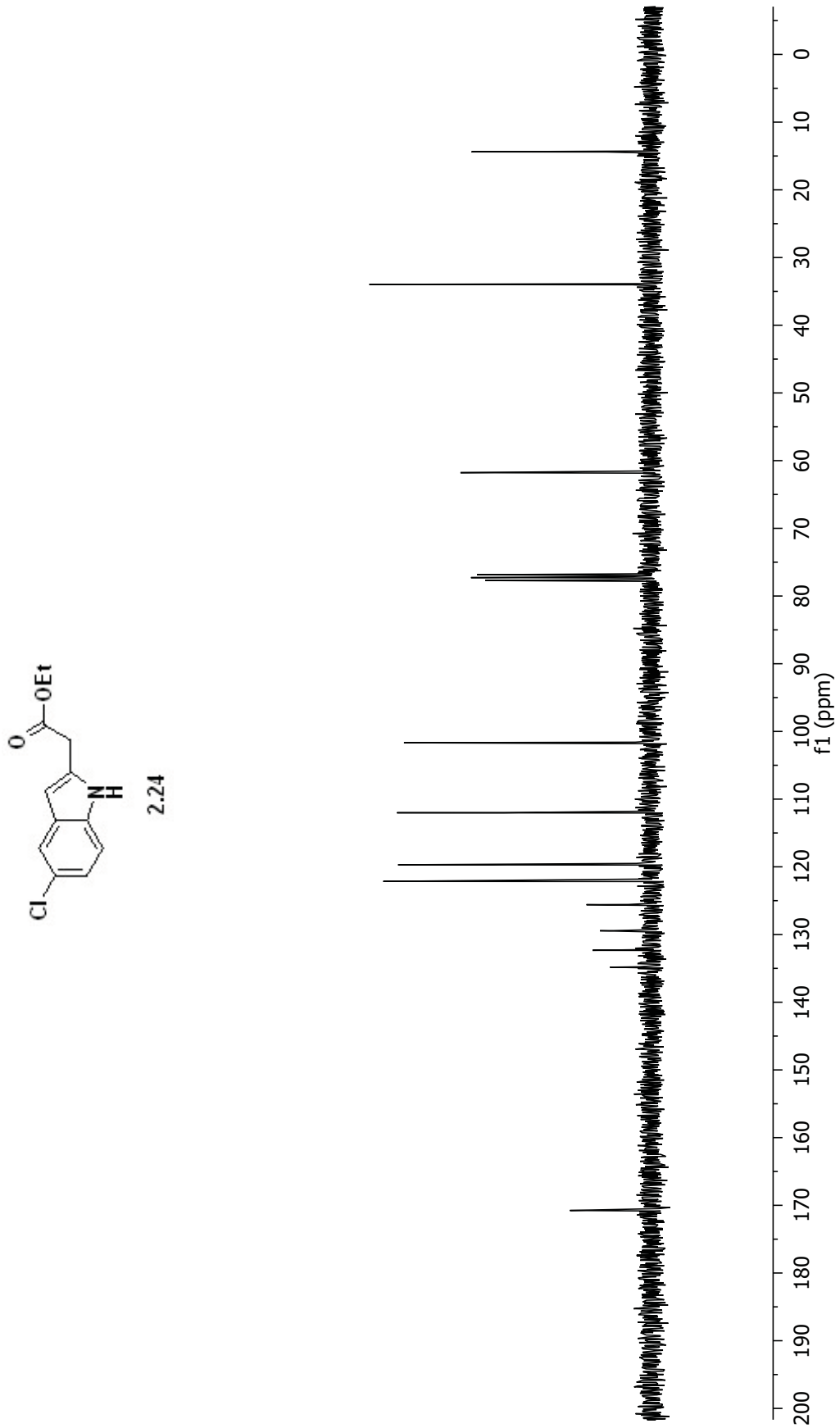

COOtBu
MeO
O
N
H
O
NH2•HCl
COOtBu
2.18
0.86
0.92
2.66
0.96
0.87
0.87
1.65
0.93
0.87
9.17
8.55
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
f1 (ppm)

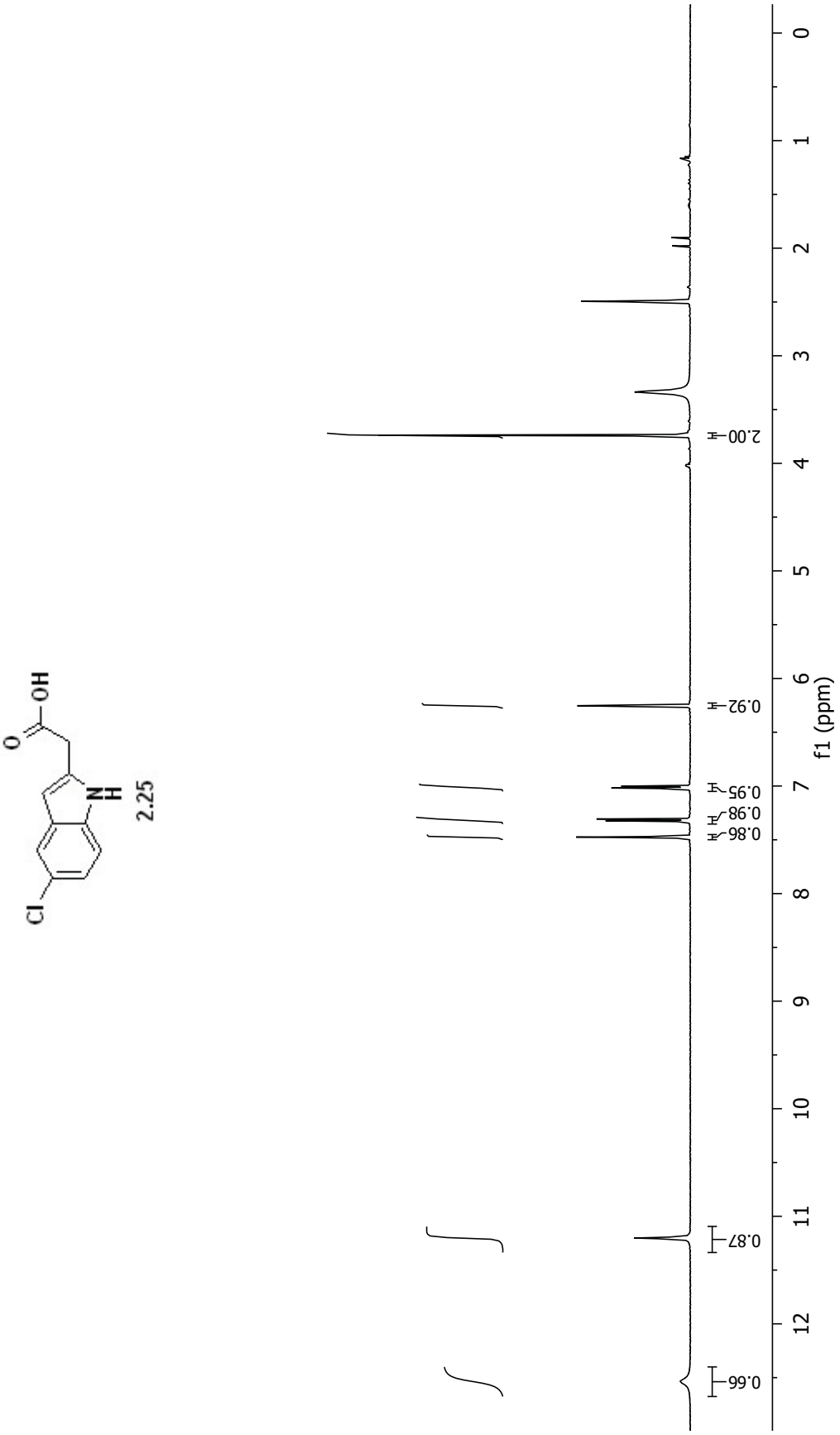

COOtBu
MeO
O
N
H
O
NH2•HCl
COOtBu

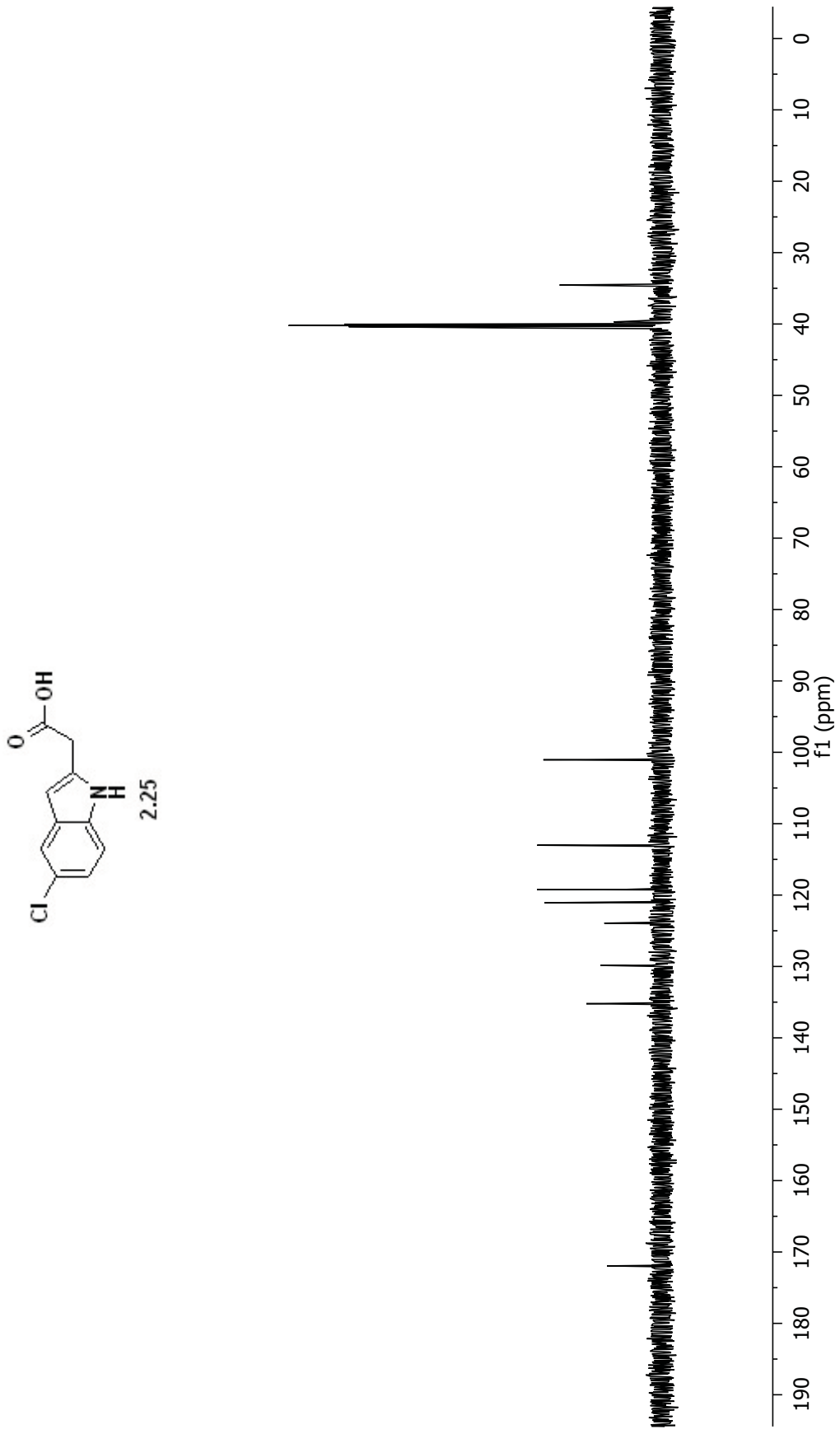
2.18

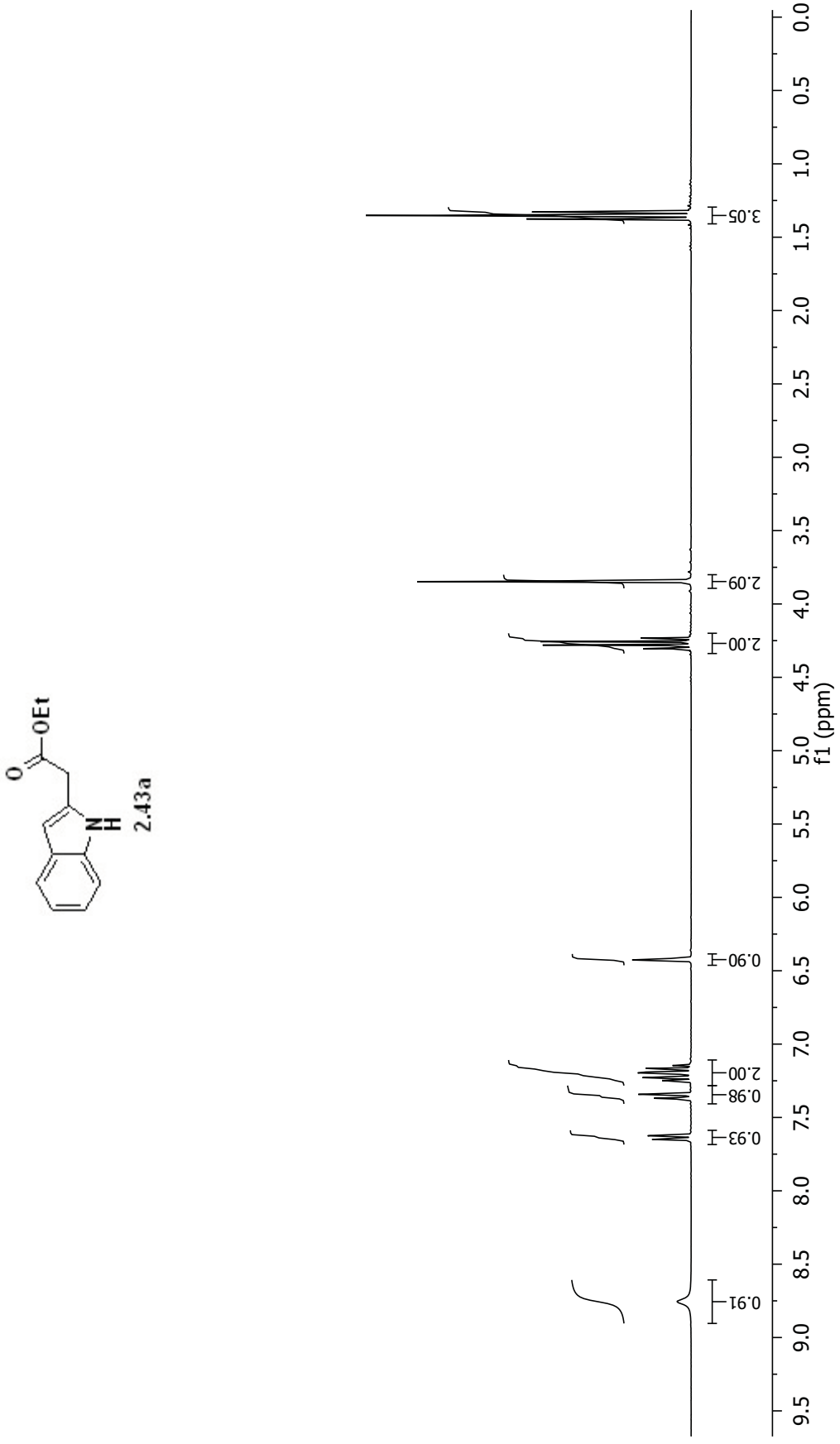

200 190 180 170 160 150 140 130 120 110 100 90 80 70 60 50 40 30 20 10 0 -10
f1 (ppm)

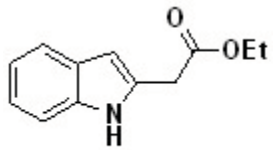

OEt
O
N
H
Cl
2.24
3.15
2.18
2.05
1.01
1.01
1.10
0.92
0.59
10.0
9.5
9.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
f1 (ppm)


O
OEt
N
H
Cl
2.24
f1 (ppm)
200
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0

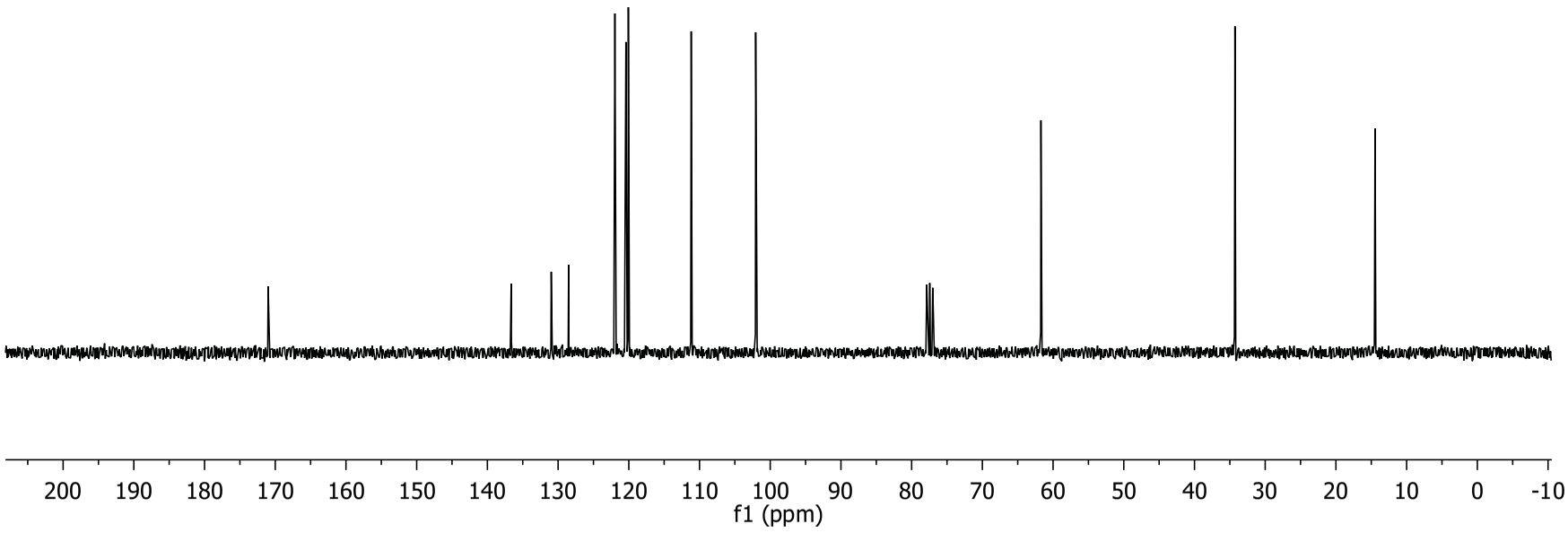

O
OH
N
H
Cl
2.25
2.00
0.92
0.95
0.98
0.86
0.87
0.66
f1 (ppm)

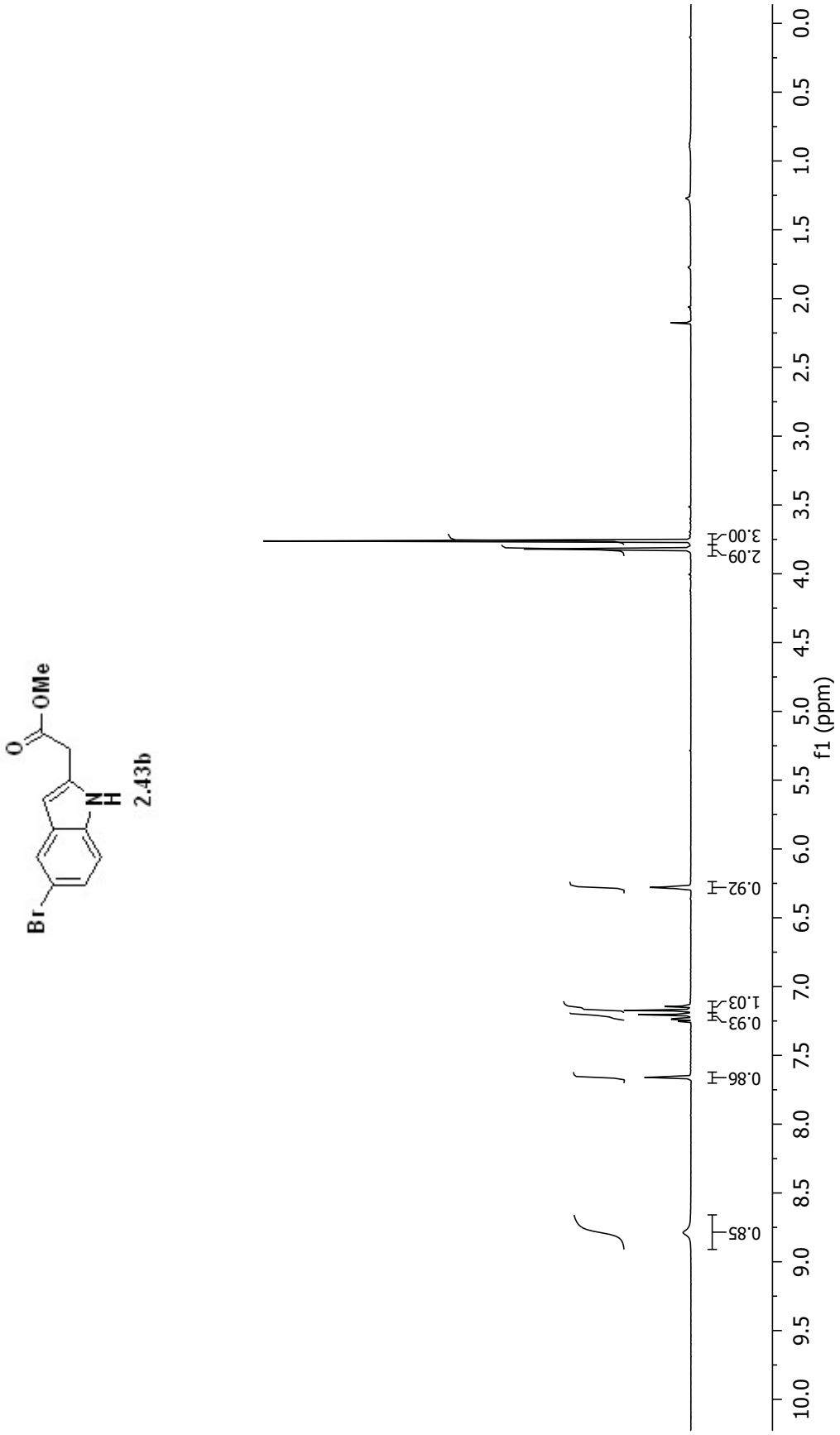

O
OH
N
H
Cl
2.25
0
10
20
30
40
50
60
70
80
90
100
110
120
130
140
150
160
170
180
190
f1 (ppm)

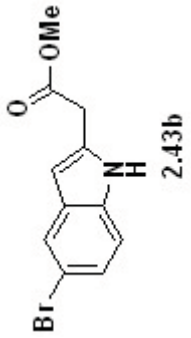

O
OEt
N
H
2.43a
3.05
2.09
2.00
0.90
2.00
0.98
0.93
0.91
9.5
9.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
f1 (ppm)


O
OEt
N
H

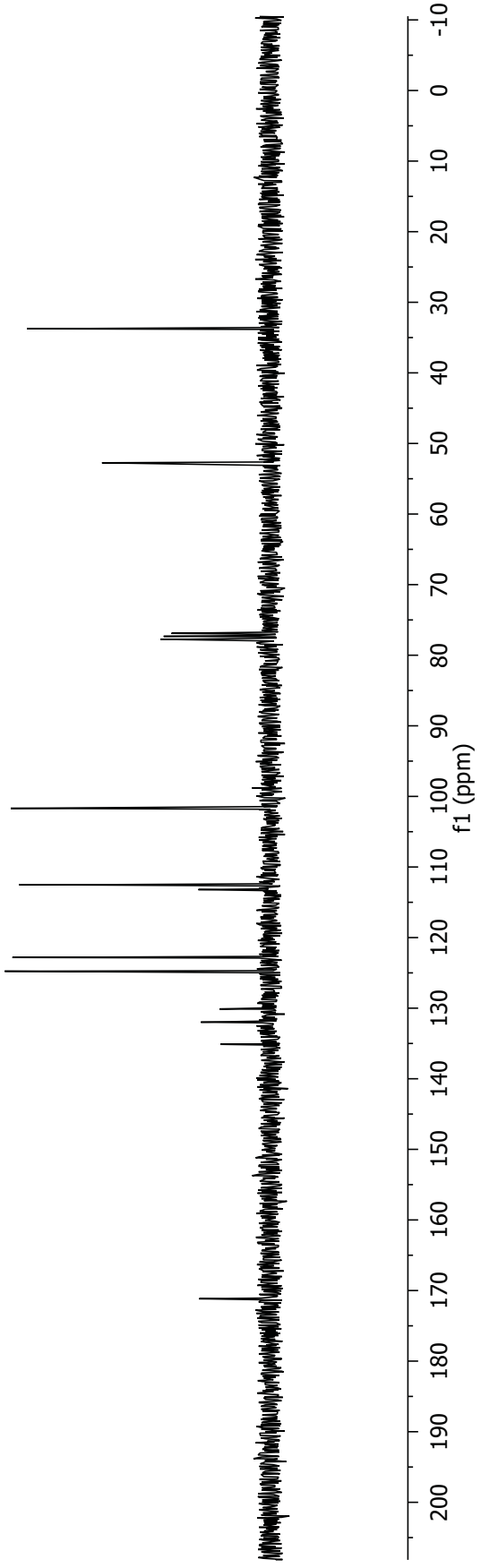
2.43a

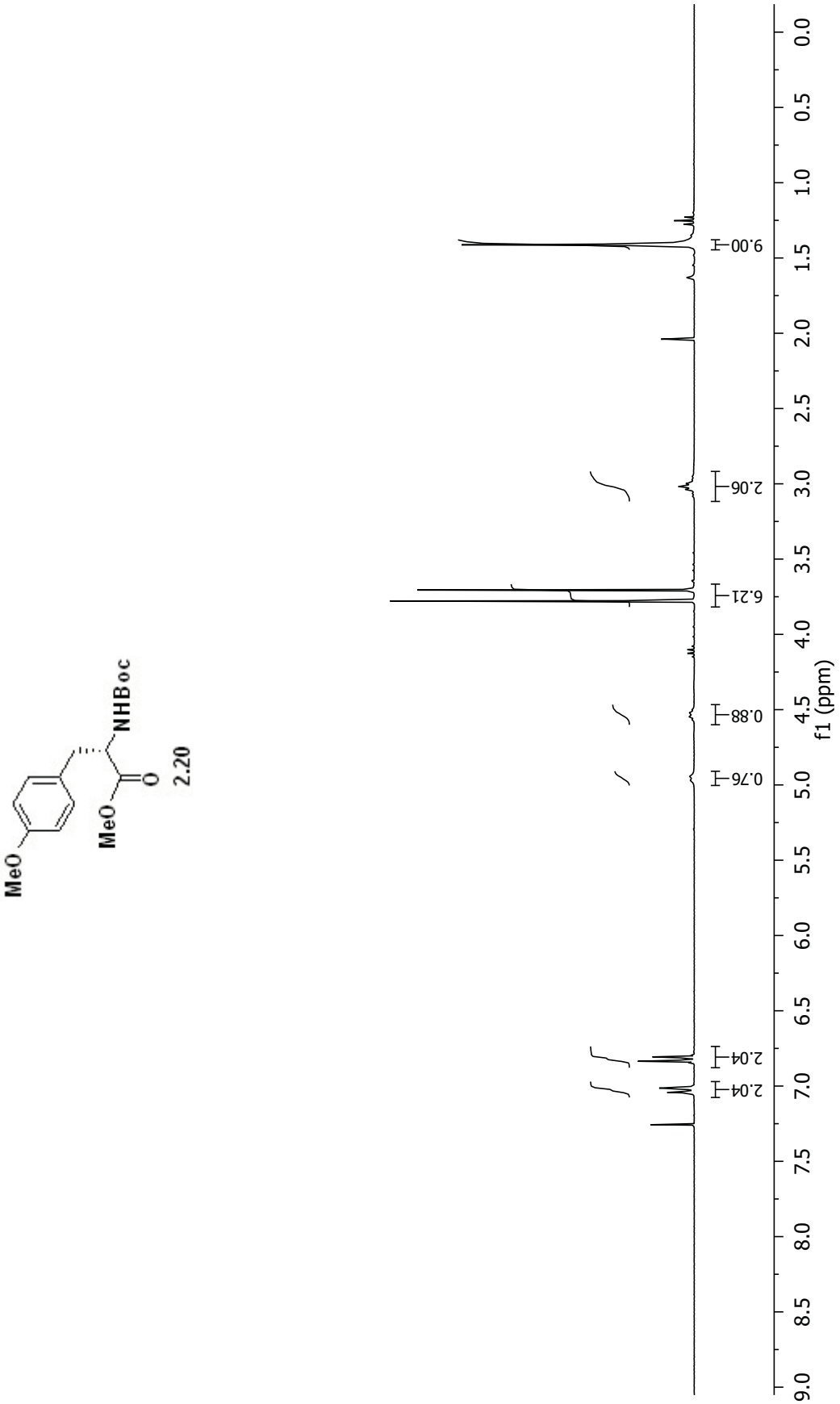

200
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
-10
f1 (ppm)

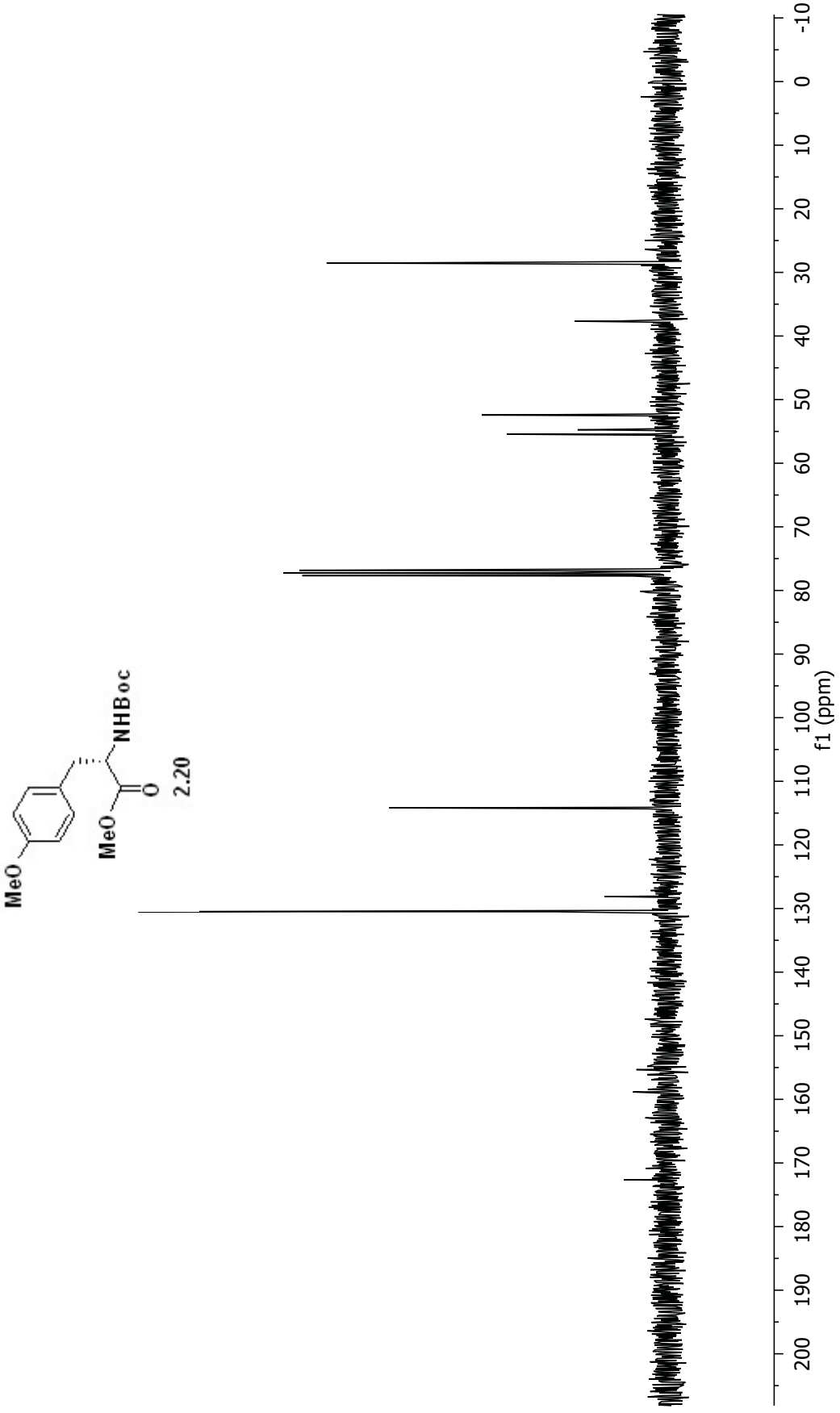

O
OMe
N
H
Br
2.43b
0.85
0.86
0.93
1.03
0.92
2.09
3.00
10.0
9.5
9.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
f1 (ppm)

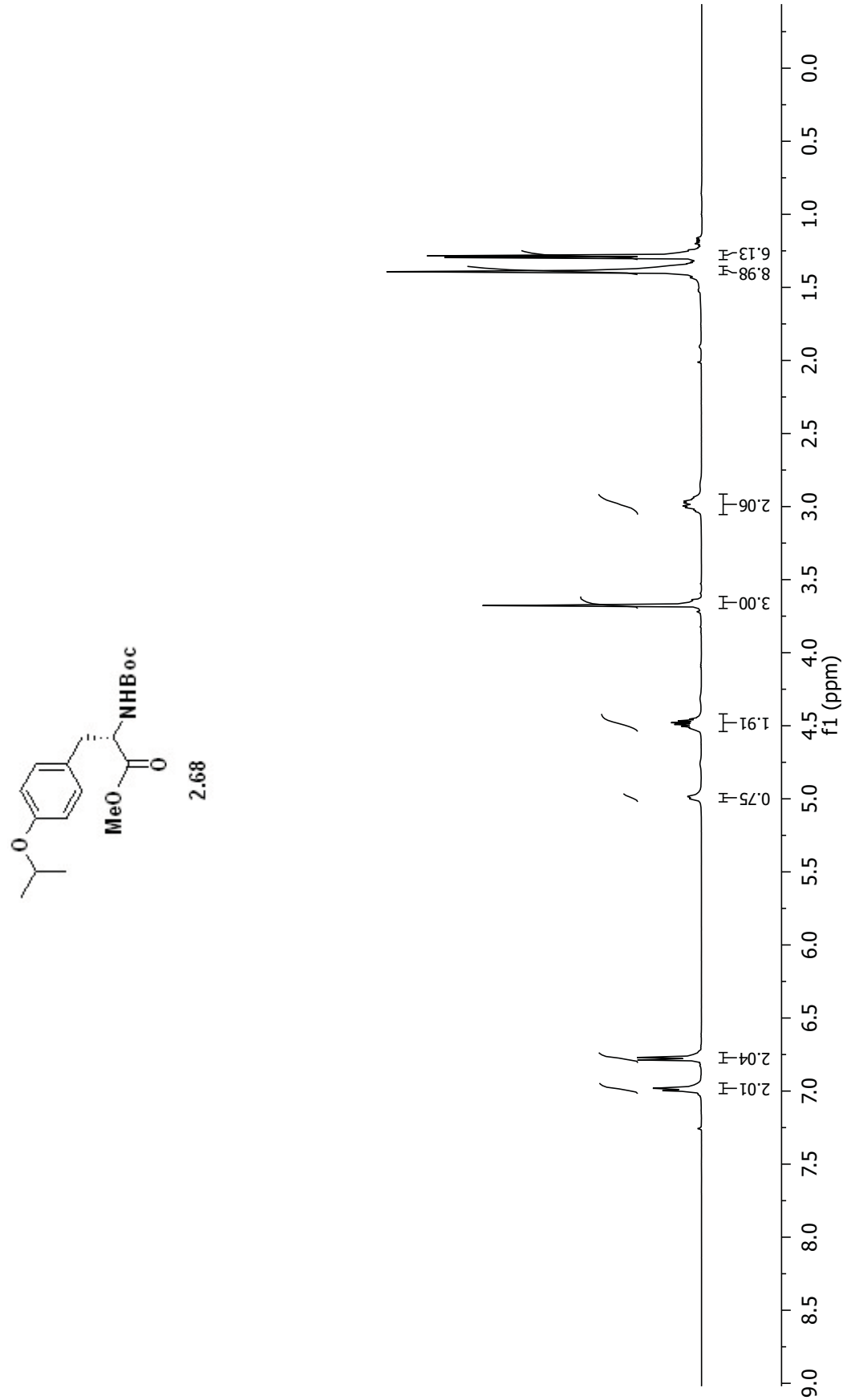

O
OMe
N
H
Br

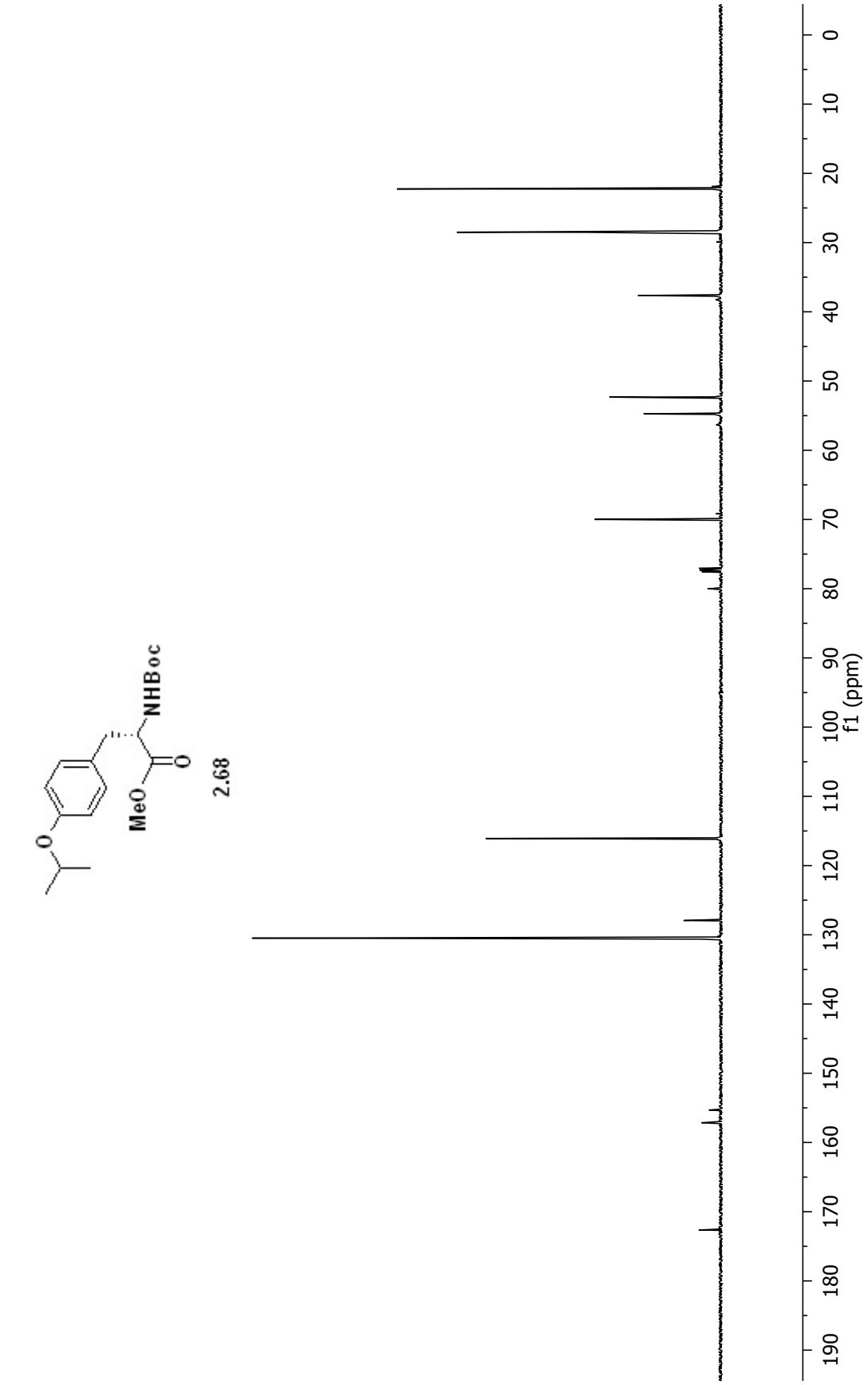
2.43b

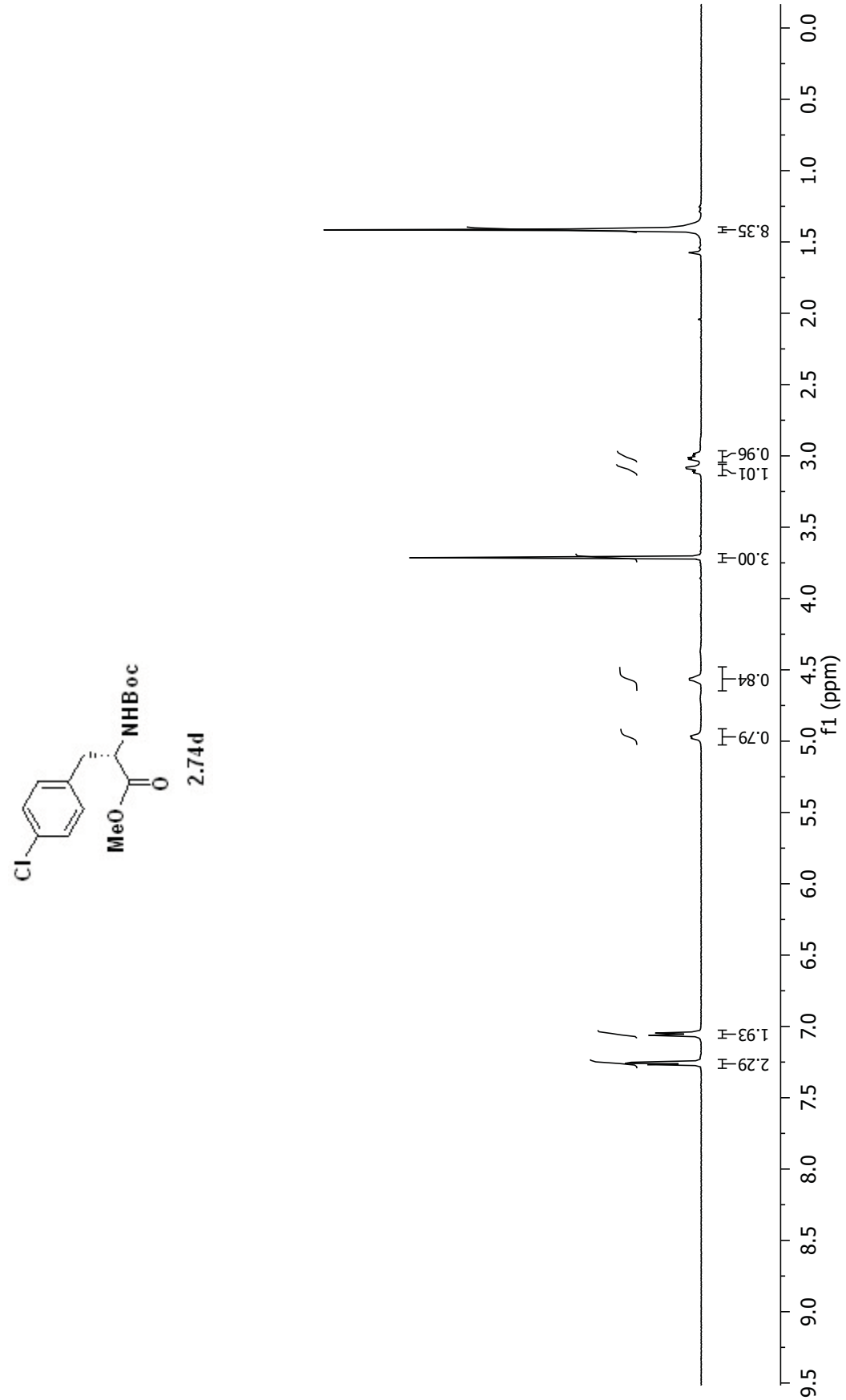

200
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
-10
f1 (ppm)

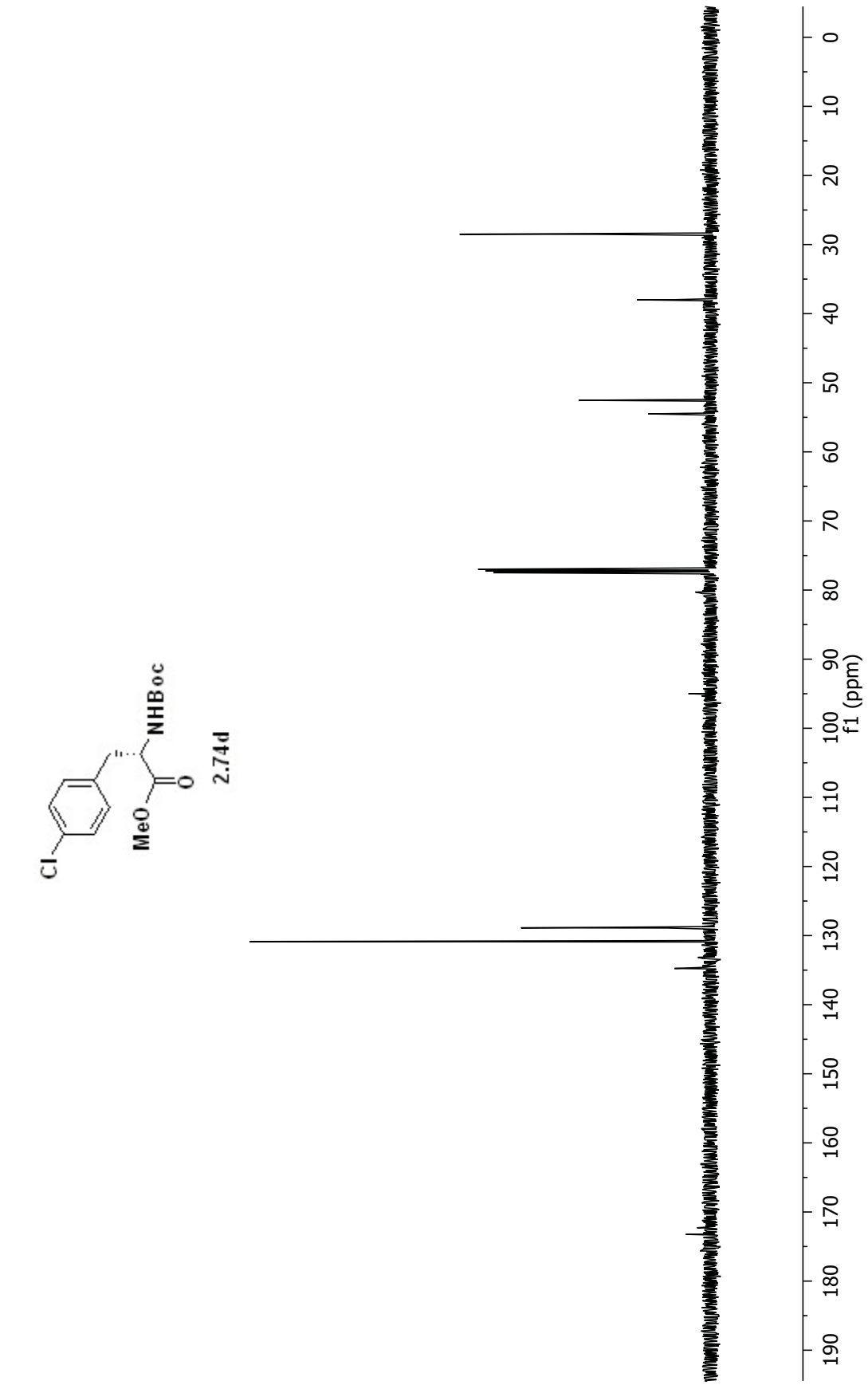

MeO
MeO
O
NHBoc
2.20
9.00
2.06
6.21
0.88
0.76
2.04
2.04
9.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
f1 (ppm)

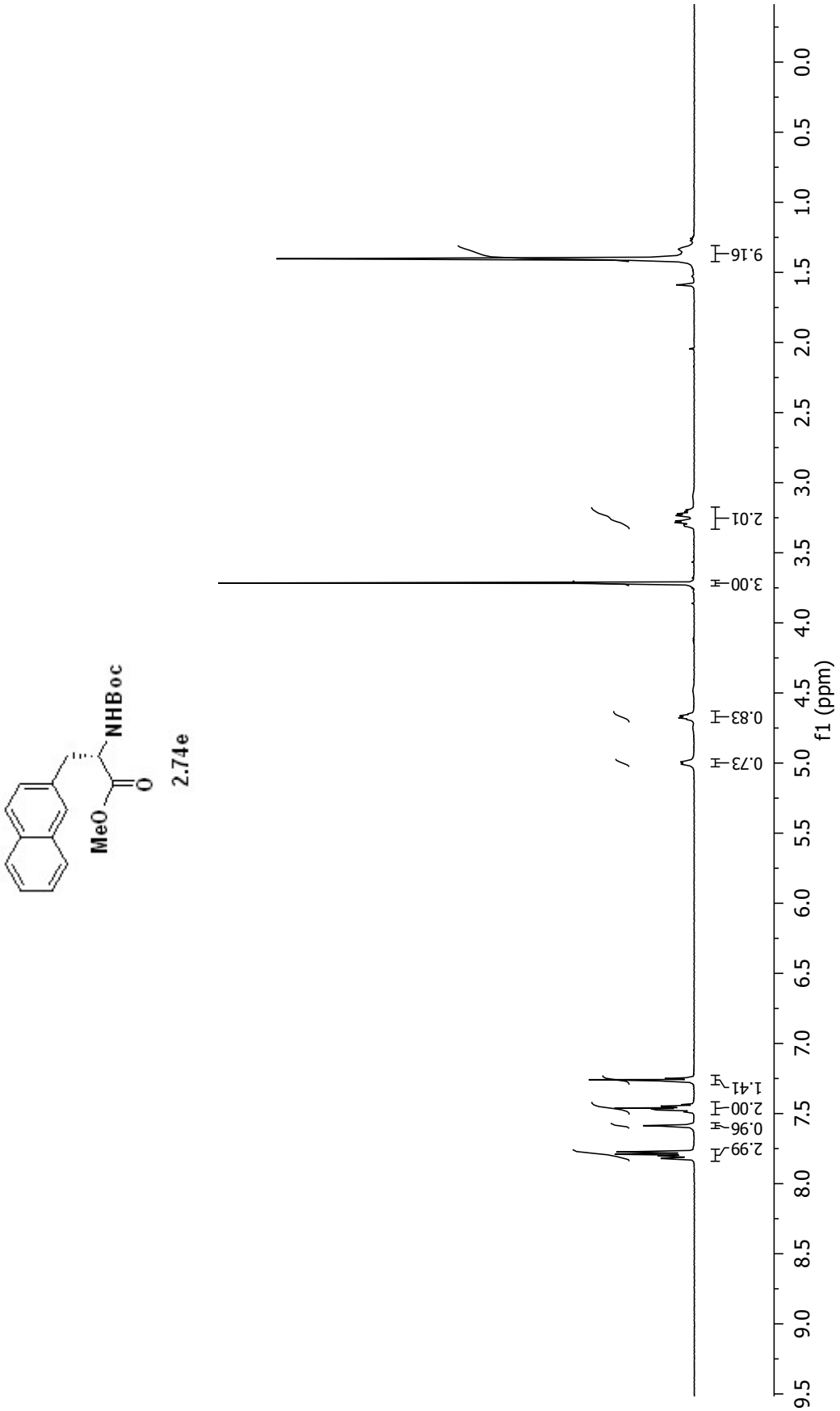

MeO
MeO
O
NHBoc
2.20
200
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
-10
f1 (ppm)

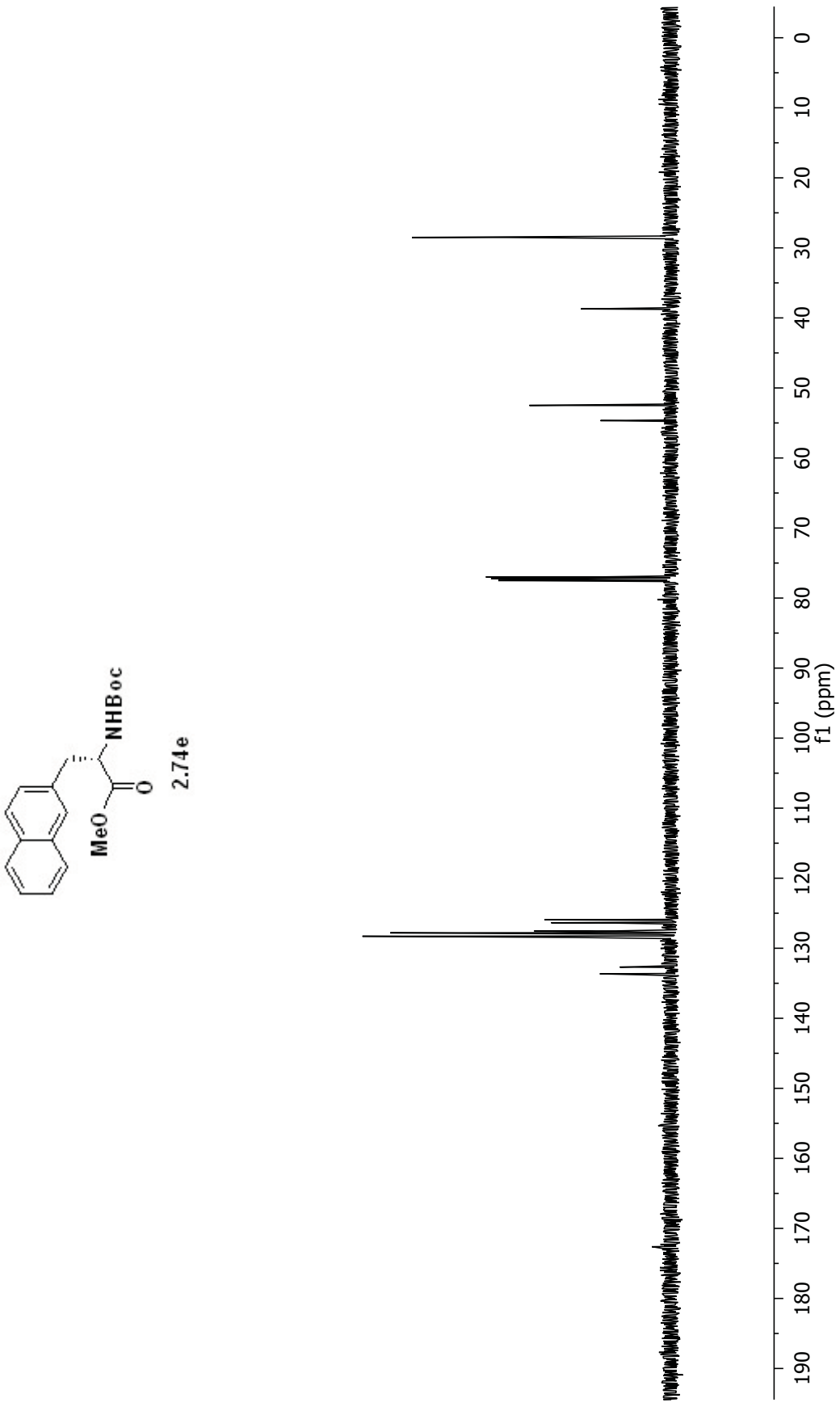

NHBoc
MeO
O
O
2.68
9.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
f1 (ppm)
2.01
2.04
0.75
1.91
3.00
2.06
8.98
6.13

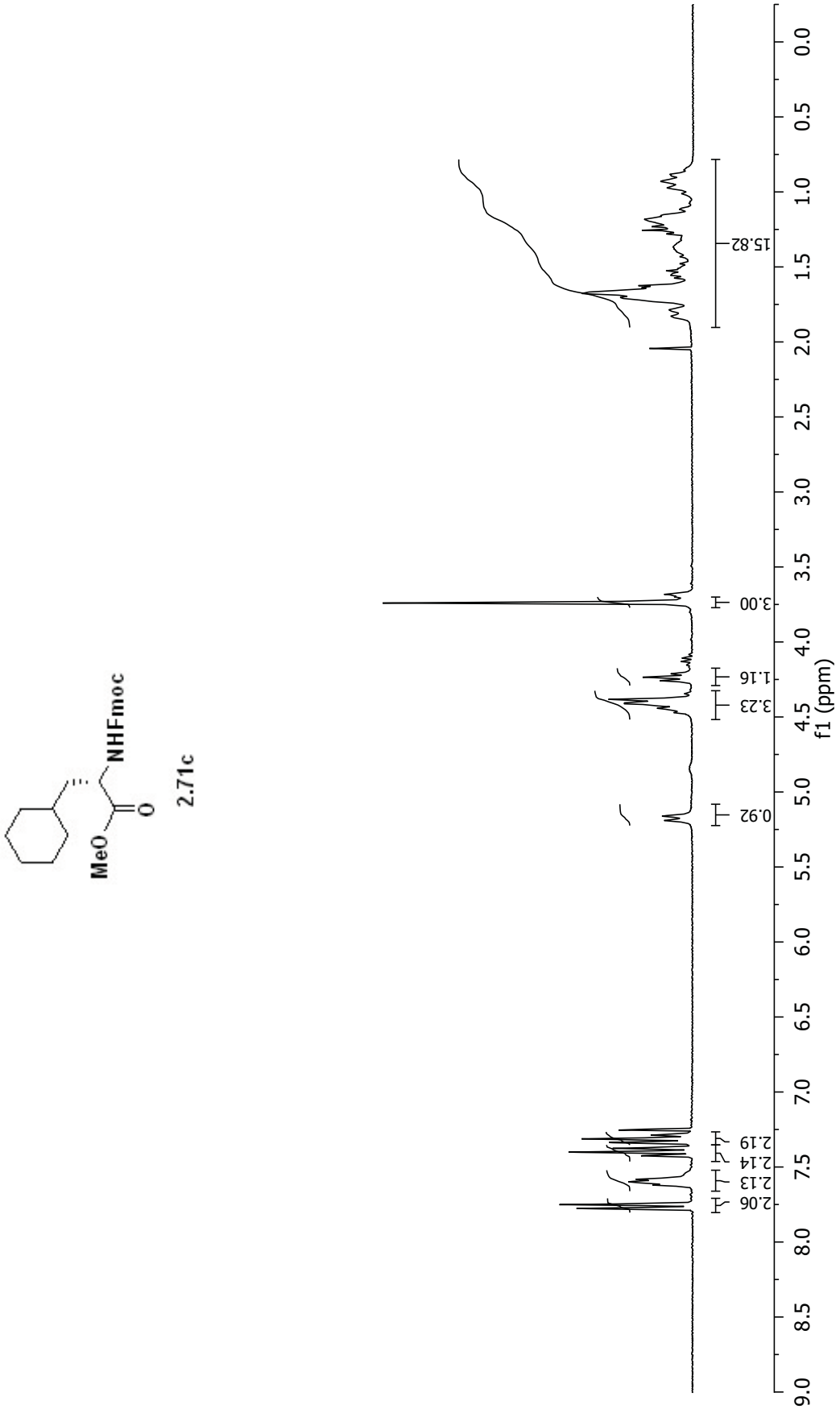

NHBoc
MeO
O
O
2.68
f1 (ppm)

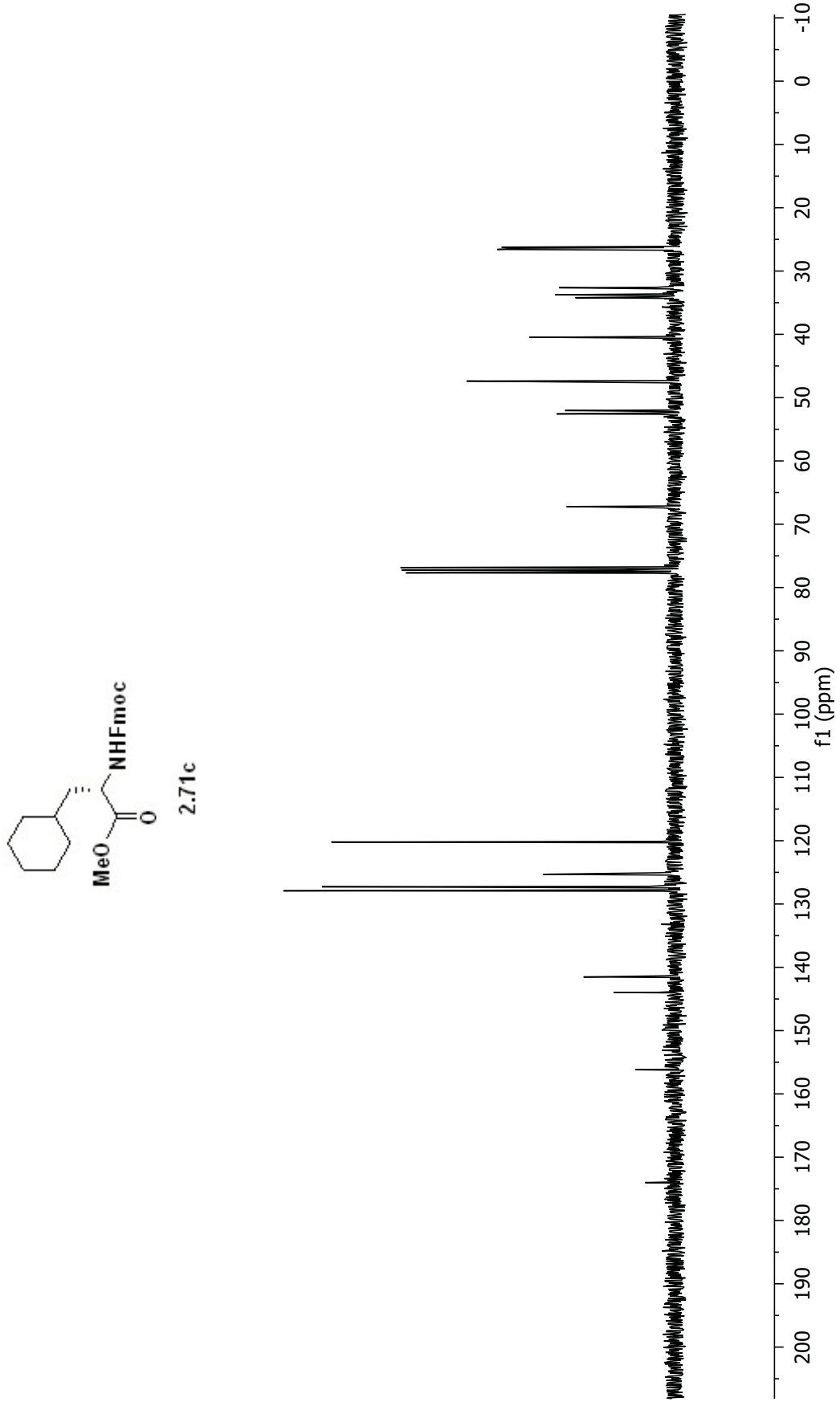

Cl
MeO
O
NHBoc
2.74d
8.35
3.00
1.01
0.96
0.84
0.79
1.93
2.29
f1 (ppm)

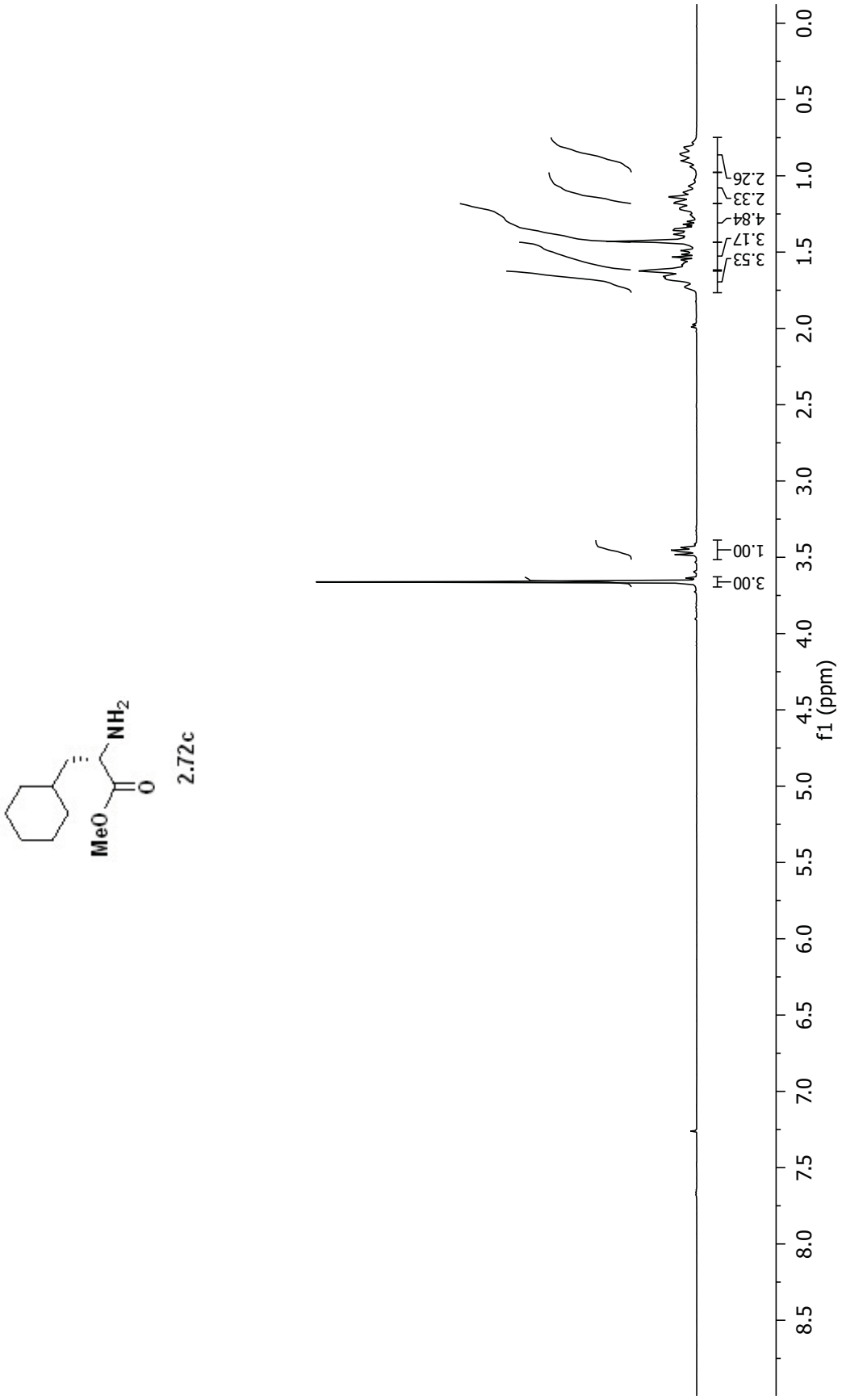

Cl
MeO
O
NHBoc
2.74d
f1 (ppm)

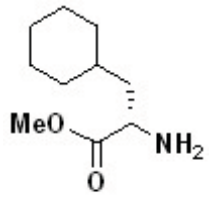

NHBoc
MeO
O
2.74e
9.16
2.01
3.00
0.83
0.73
2.99
0.96
2.00
1.41
9.5
9.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
f1 (ppm)


NHBoc
MeO
O
2.74e
f1 (ppm)

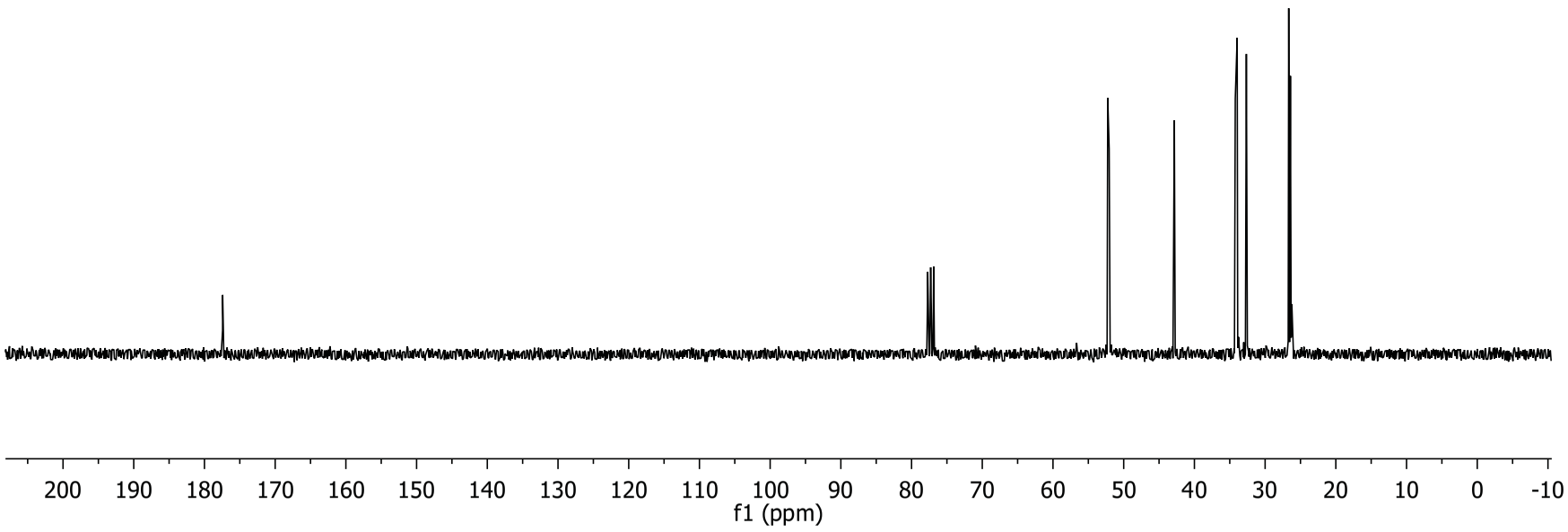

NHFmoc
MeO
O
2.71c
2.06
2.13
2.14
2.19
0.92
3.23
1.16
3.00
15.82
f1 (ppm)

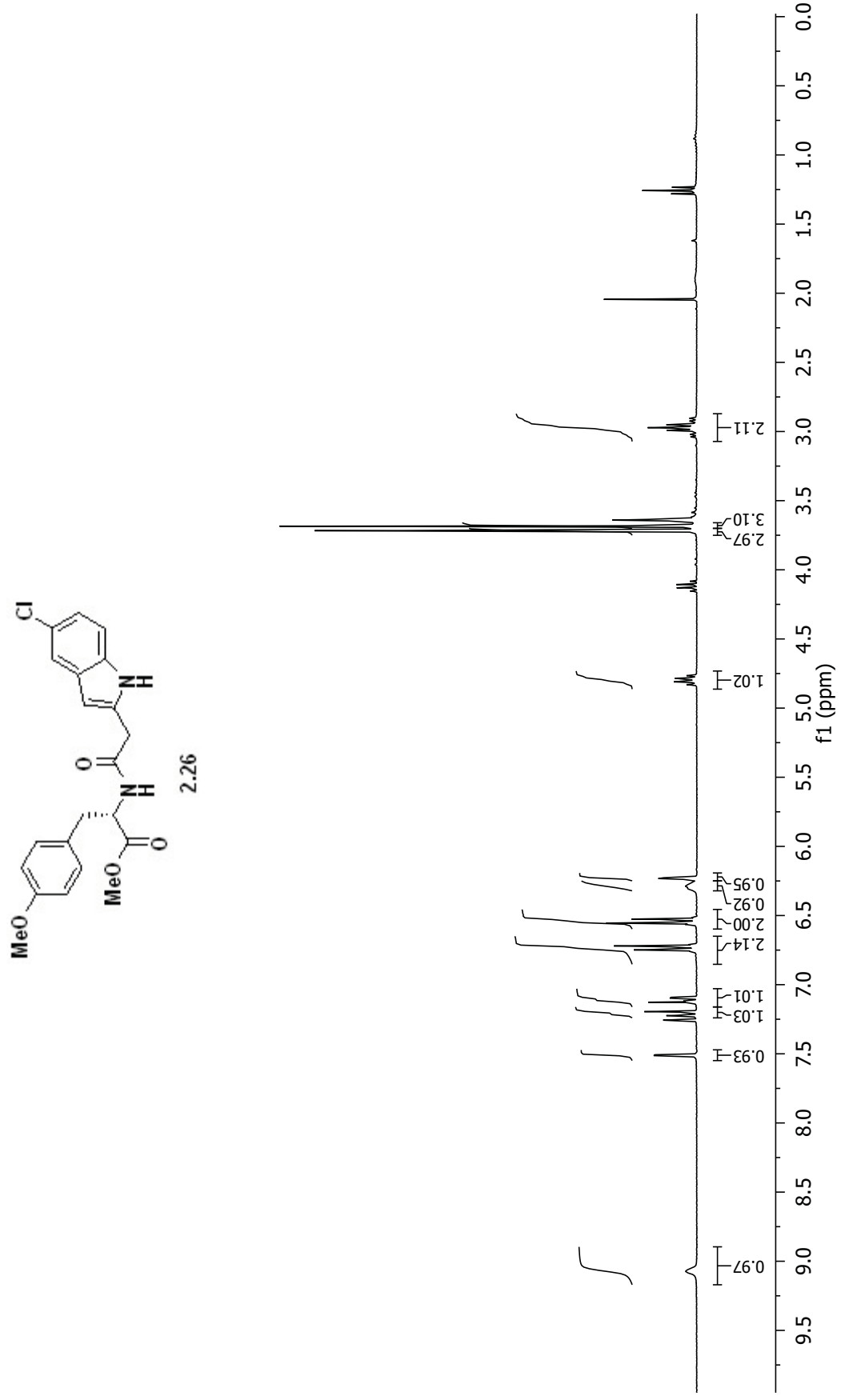

NHFmoc
MeO
O
2.71c
f1 (ppm)
200
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
-10

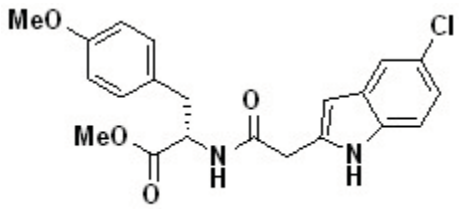

NH2
MeO
O
2.72c
3.00
1.00
3.53
3.17
4.84
2.33
2.26
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
f1 (ppm)


MeO
NH2
O

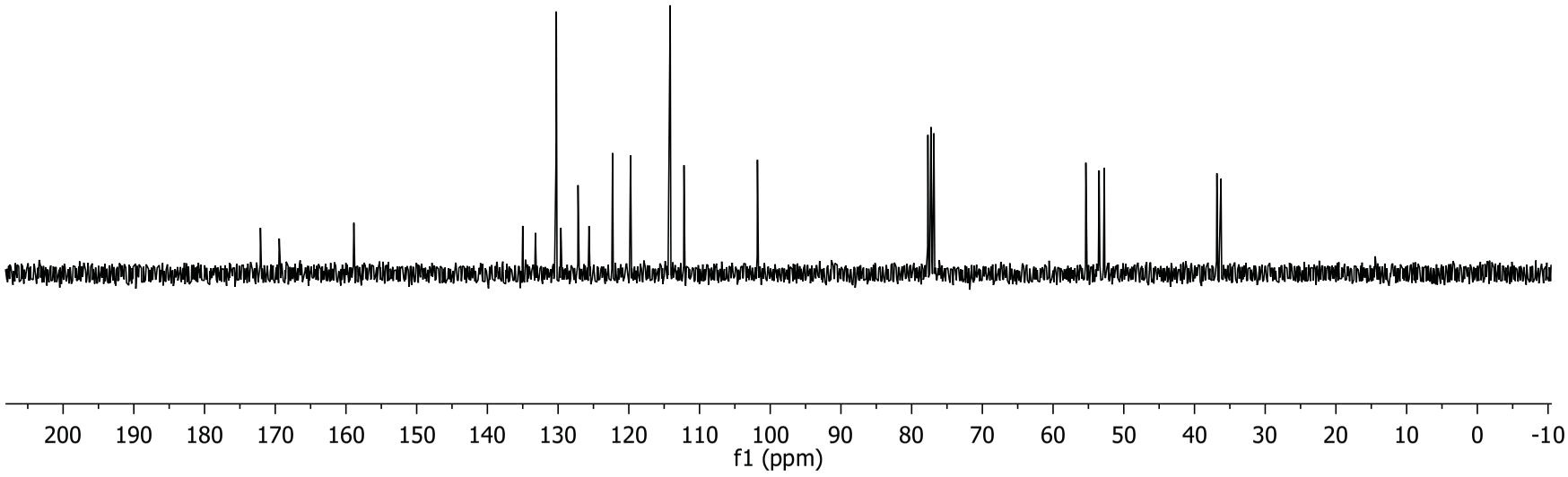
2.72c

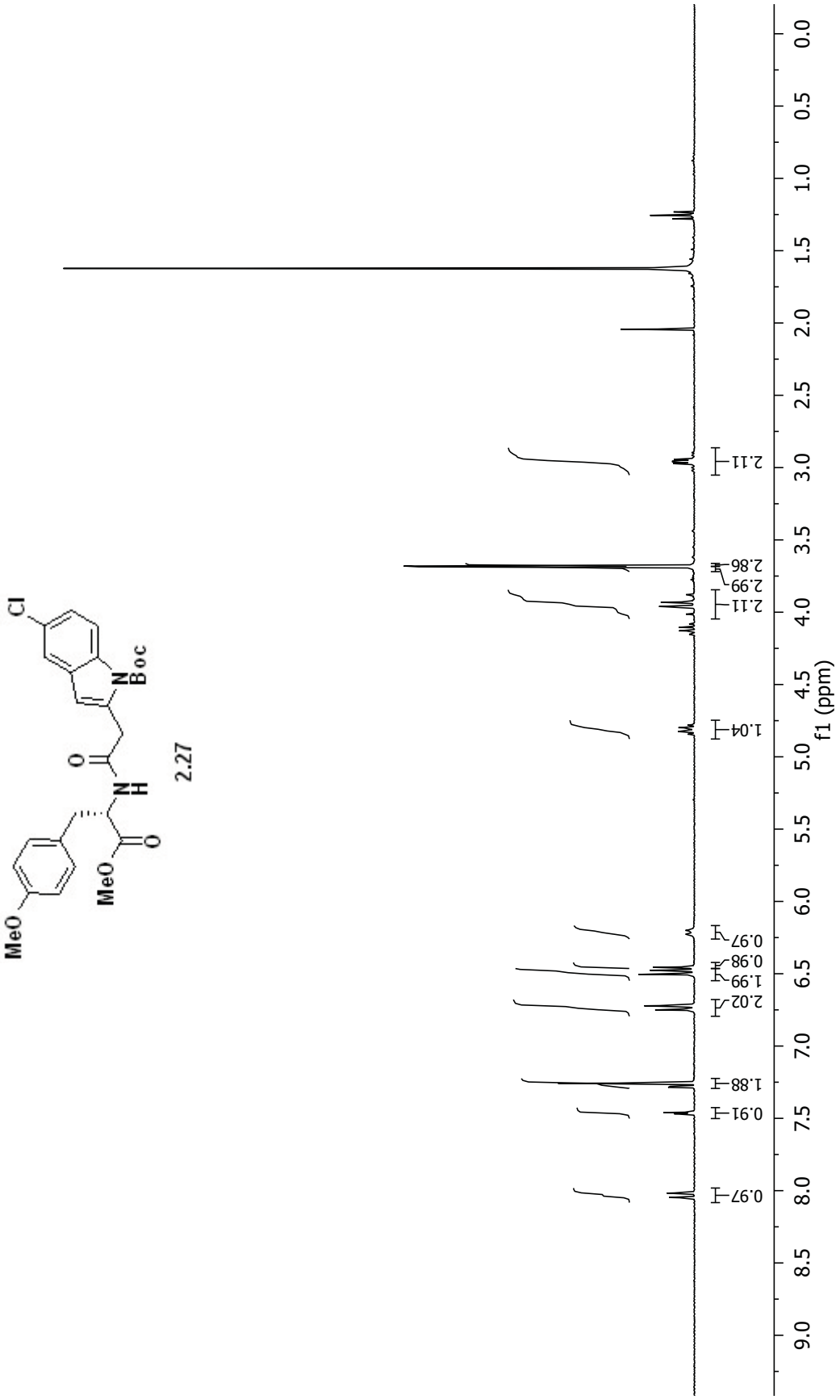

200
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
-10
f1 (ppm)

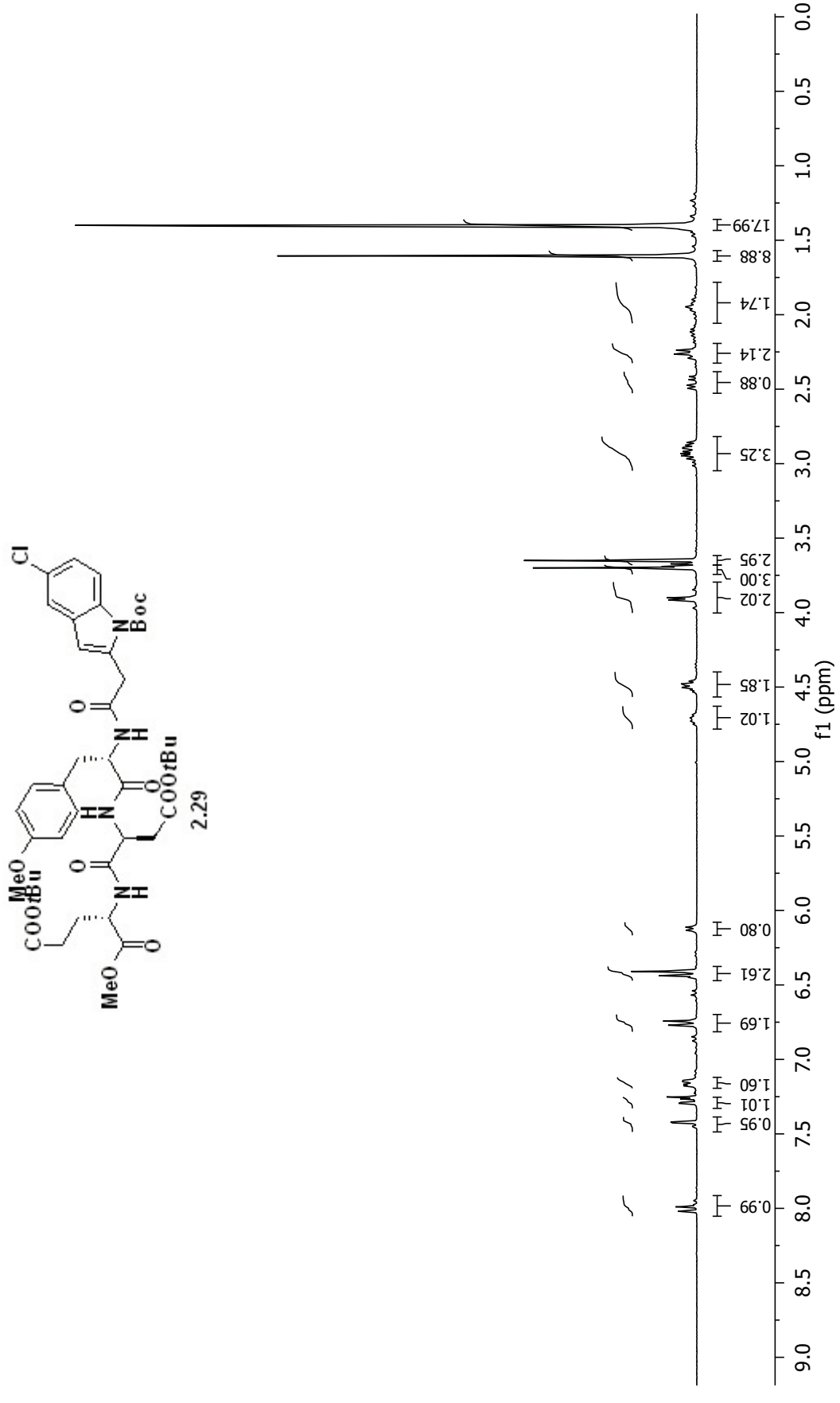

Cl
N
H
O
N
H
O
MeO
O
MeO
2.26
0.97
2.11
3.10
2.97
1.02
0.95
0.92
2.00
2.14
1.01
1.03
0.93
9.5
9.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
f1 (ppm)

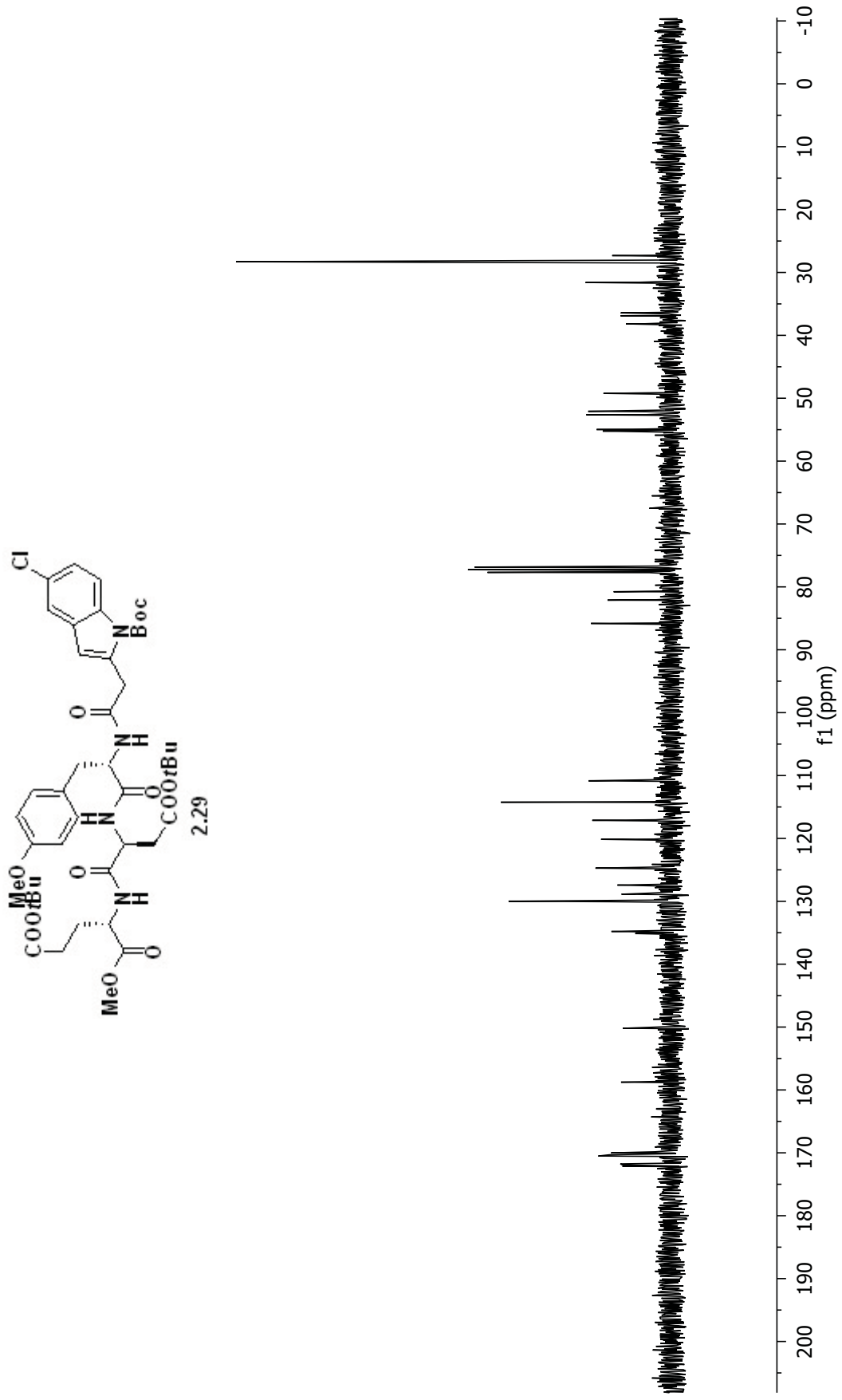

MeO
Cl
O
MeO
N
H
N
H
O

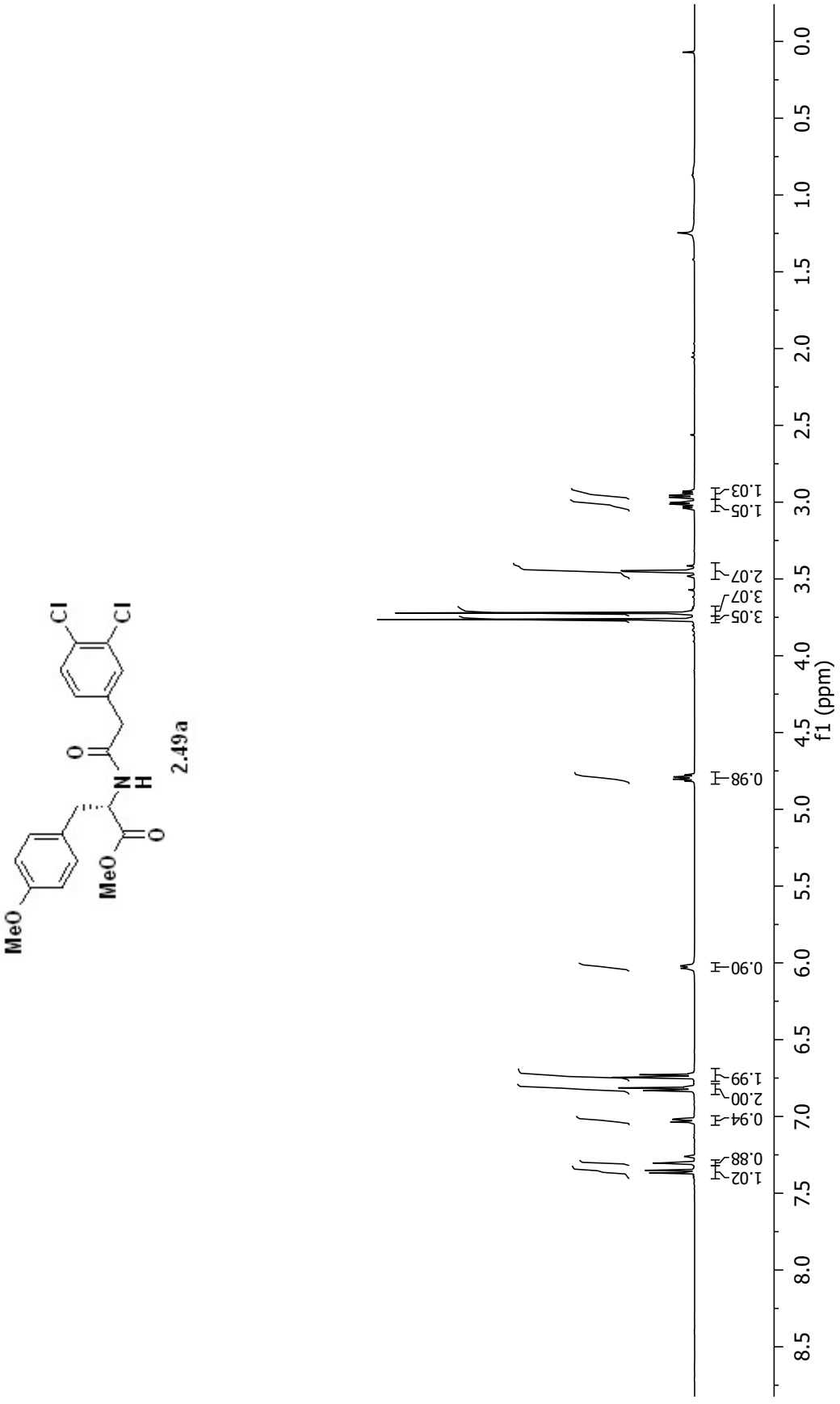
2.26

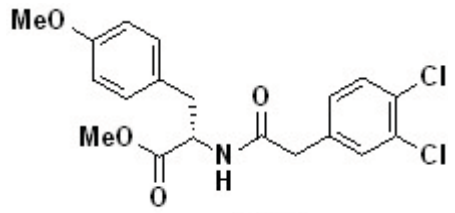

200
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
-10
f1 (ppm)


Cl
N
Boc
O
N
H
MeO
O
O
MeO
2.27
2.11
2.86
2.99
2.11
1.04
0.97
1.99
0.98
2.02
1.88
0.91
0.97
9.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
f1 (ppm)

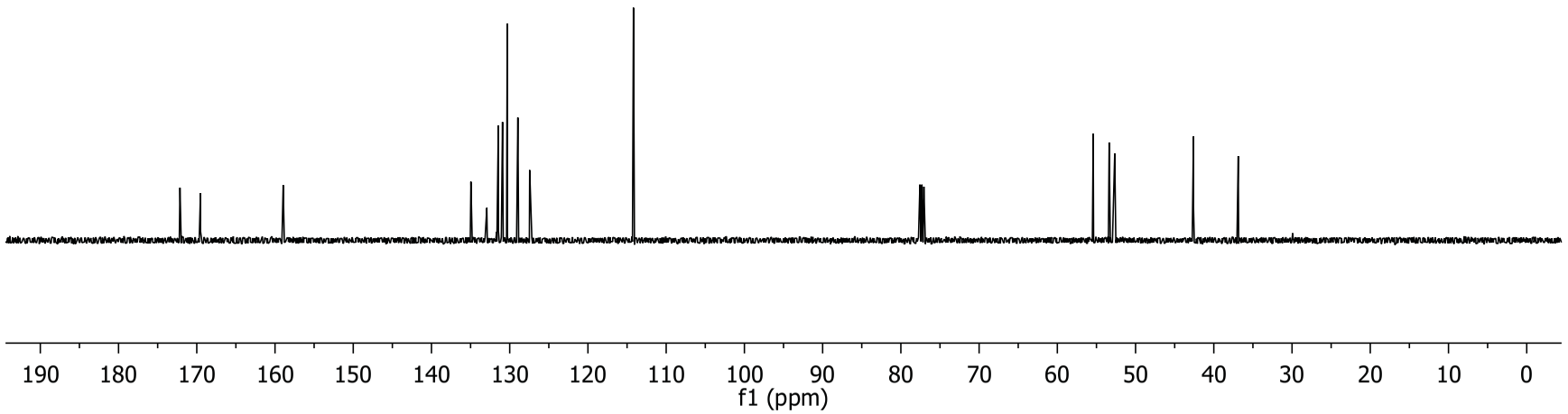

Cl
N
Boc
O
N
H
O
N
H
O
COOtBu
MeO
COOtBu
O
N
H
O
MeO
2.29
17.99
8.88
1.74
2.14
0.88
3.25
2.95
3.00
2.02
1.85
1.02
0.80
2.61
1.69
1.60
1.01
0.95
0.99
9.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
f1 (ppm)

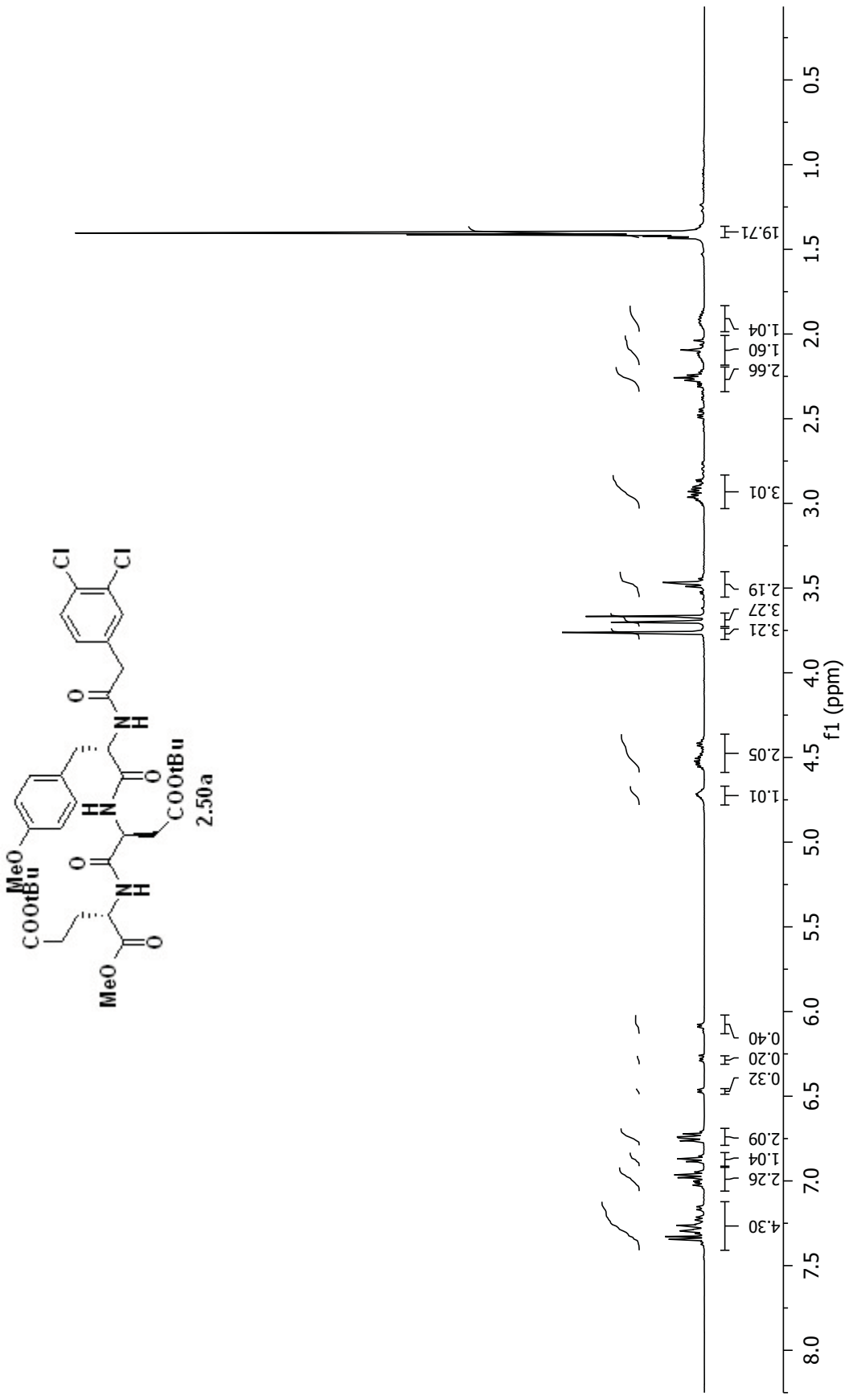

Cl
N
Boc
O
N
H
O
N
H
COOtBu
MeO
COOtBu
O
N
H
O
MeO
2.29
200
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
-10
f1 (ppm)

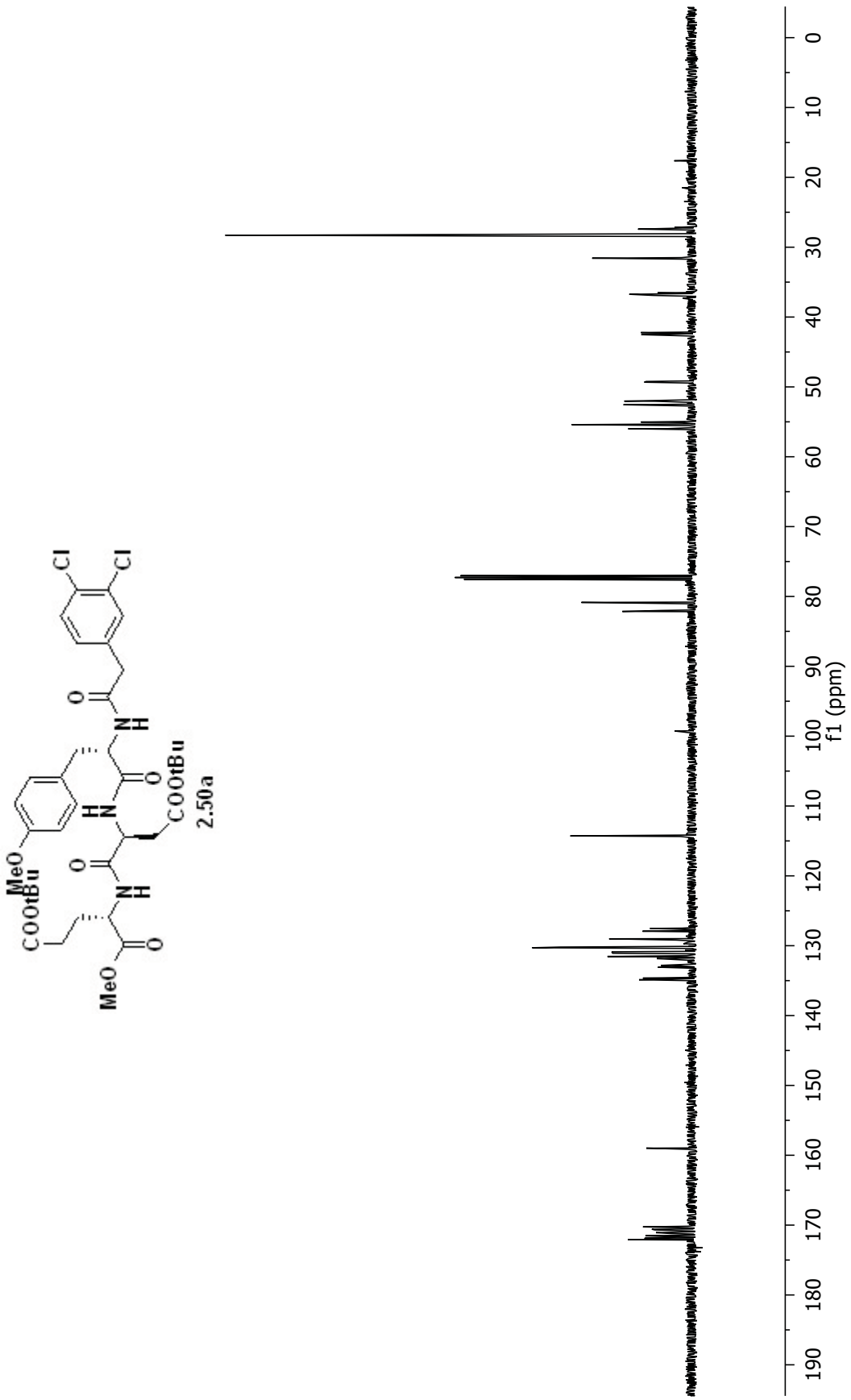

Cl
Cl
O
N
H
O
MeO
O
MeO
2.49a
1.02
0.88
0.94
2.00
1.99
0.90
0.98
3.05
3.07
2.07
1.05
1.03
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
f1 (ppm)

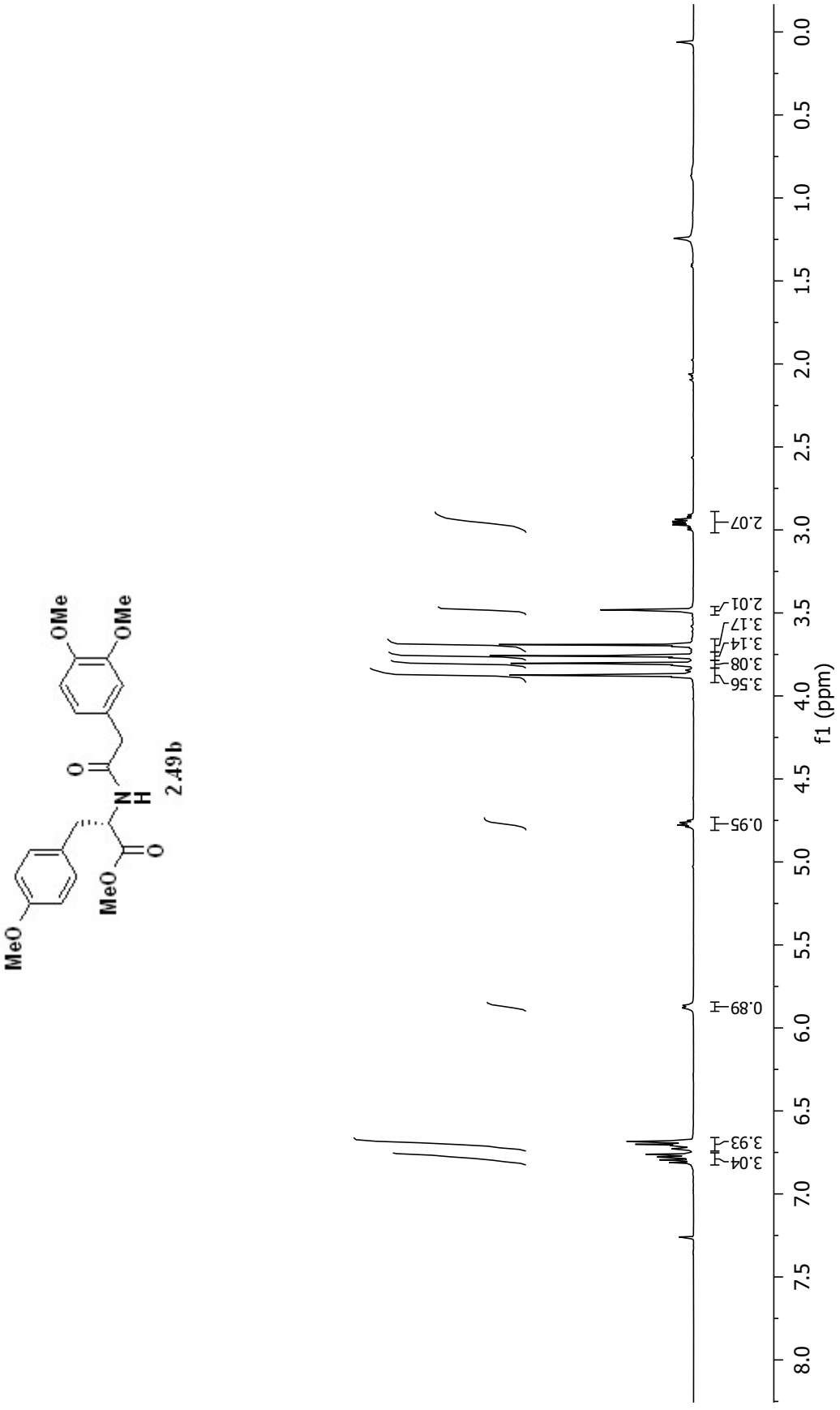

MeO
MeO
O
O
N
H
Cl
Cl

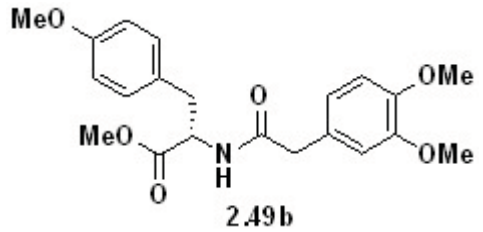
2.49a


190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
f1 (ppm)

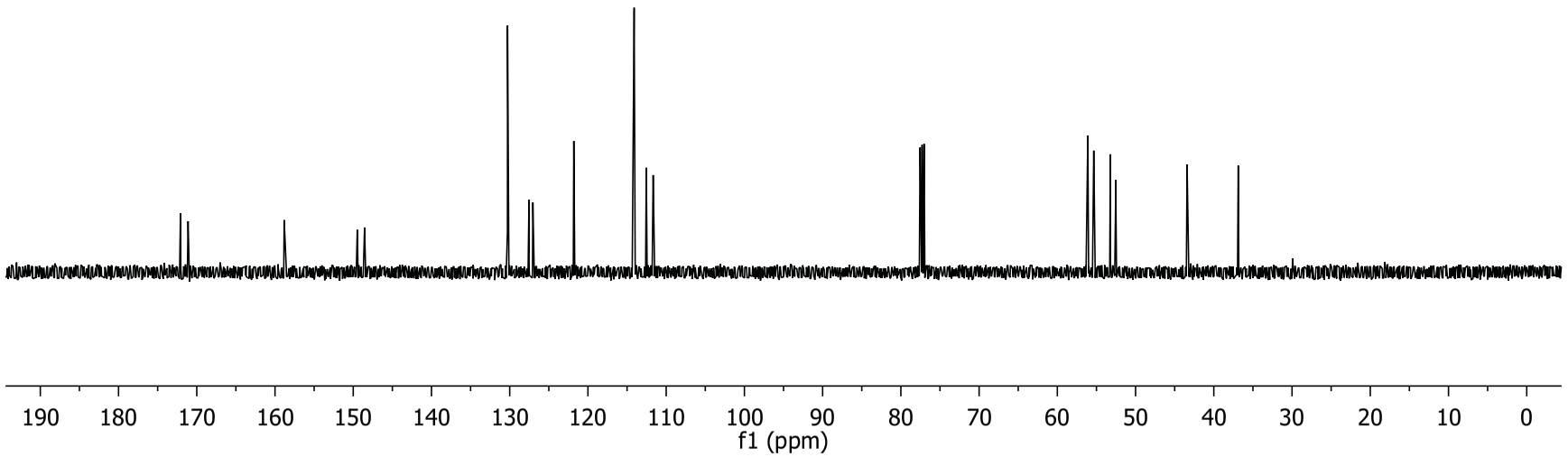

MeO
COOtBu
O
N
H
O
N
H
O
COOtBu
Cl
Cl
O
N
H
O
MeO
2.50a
19.71
4.30
2.26
1.04
2.09
0.32
0.20
0.40
1.01
2.05
3.21
3.27
2.19
3.01
2.66
1.60
1.04
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
f1 (ppm)

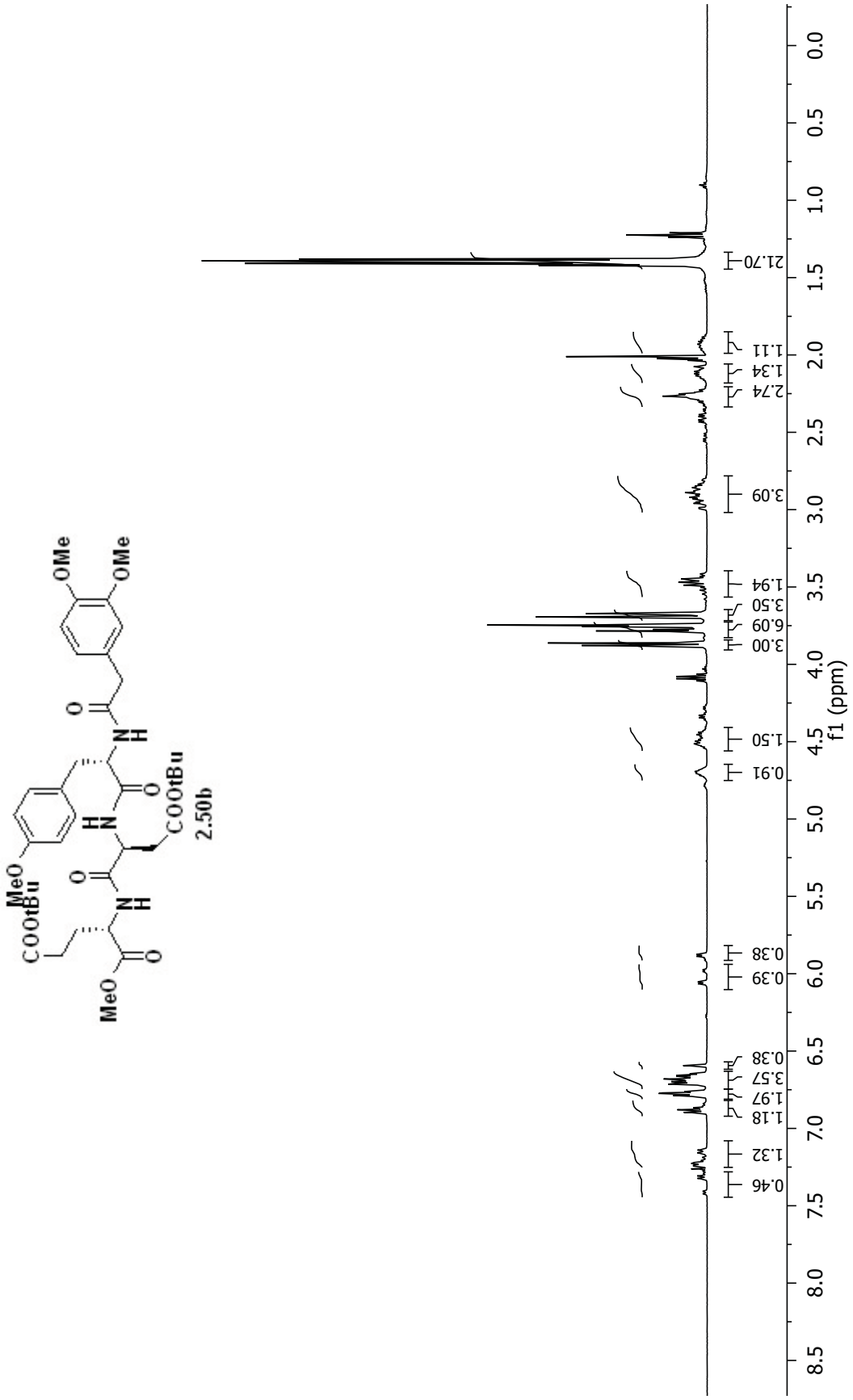

Cl
Cl
O
N
H
O
N
H
O
COOtBu
2.50a
MeO
COOtBu
O
N
H
O
MeO
f1 (ppm)
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0

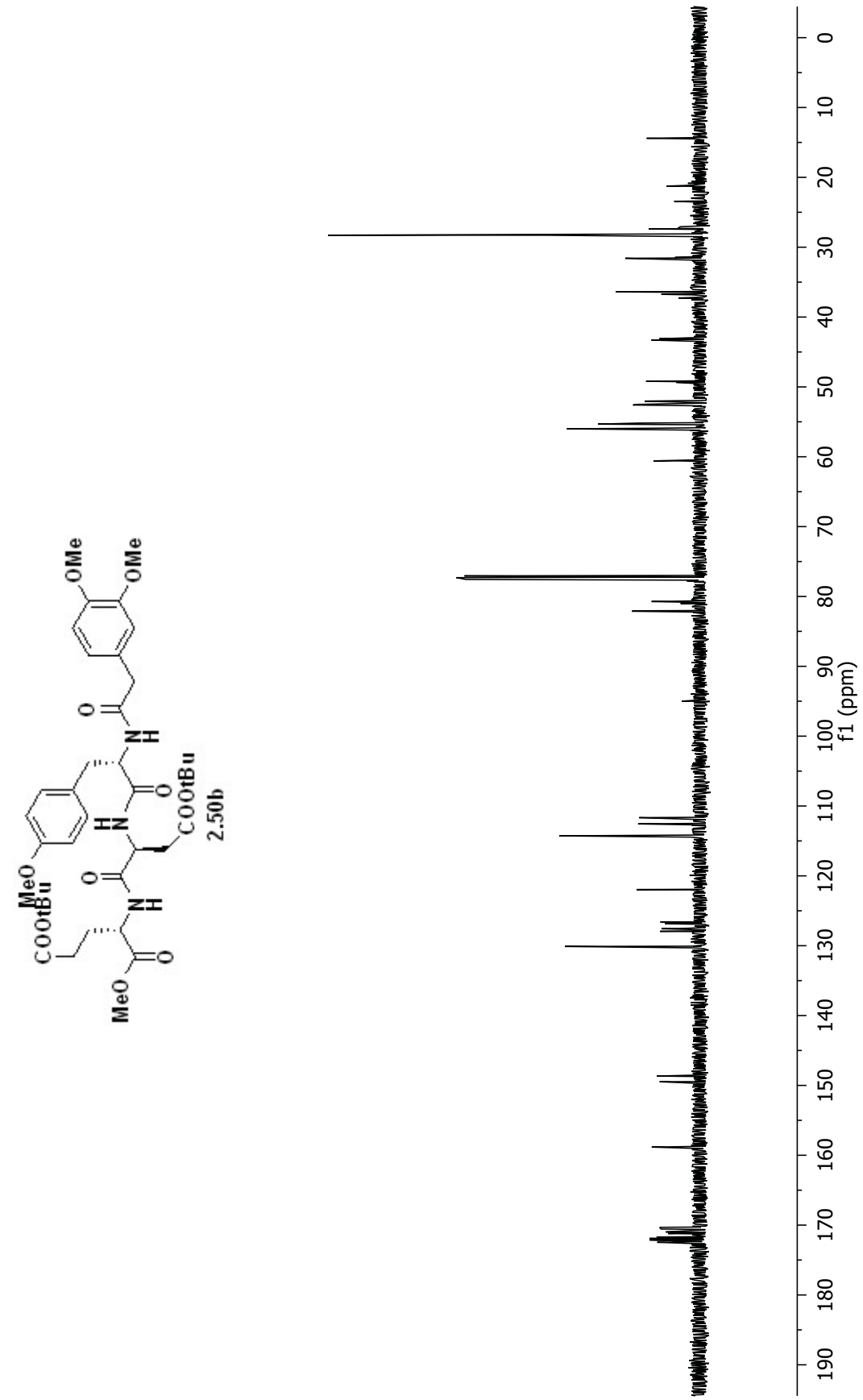

OMe
OMe
O
N
H
O
MeO
O
MeO
2.49b
3.04
3.93
0.89
0.95
3.56
3.08
3.14
3.17
2.01
2.07
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
f1 (ppm)

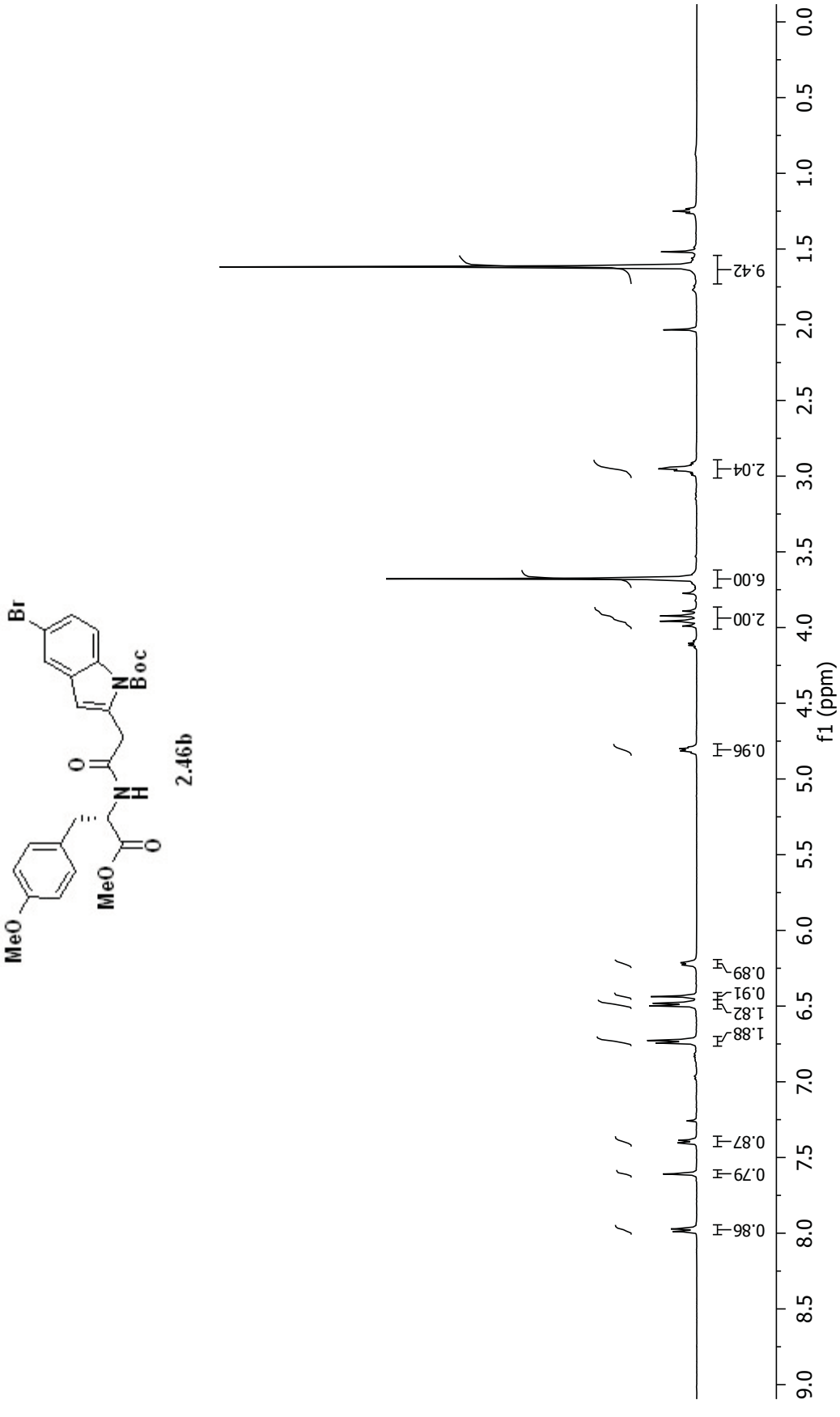

MeO
MeO
O
N
H
O
OMe
OMe

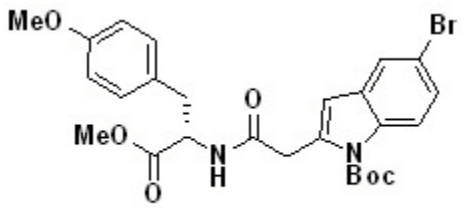
2.49b


190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
f1 (ppm)

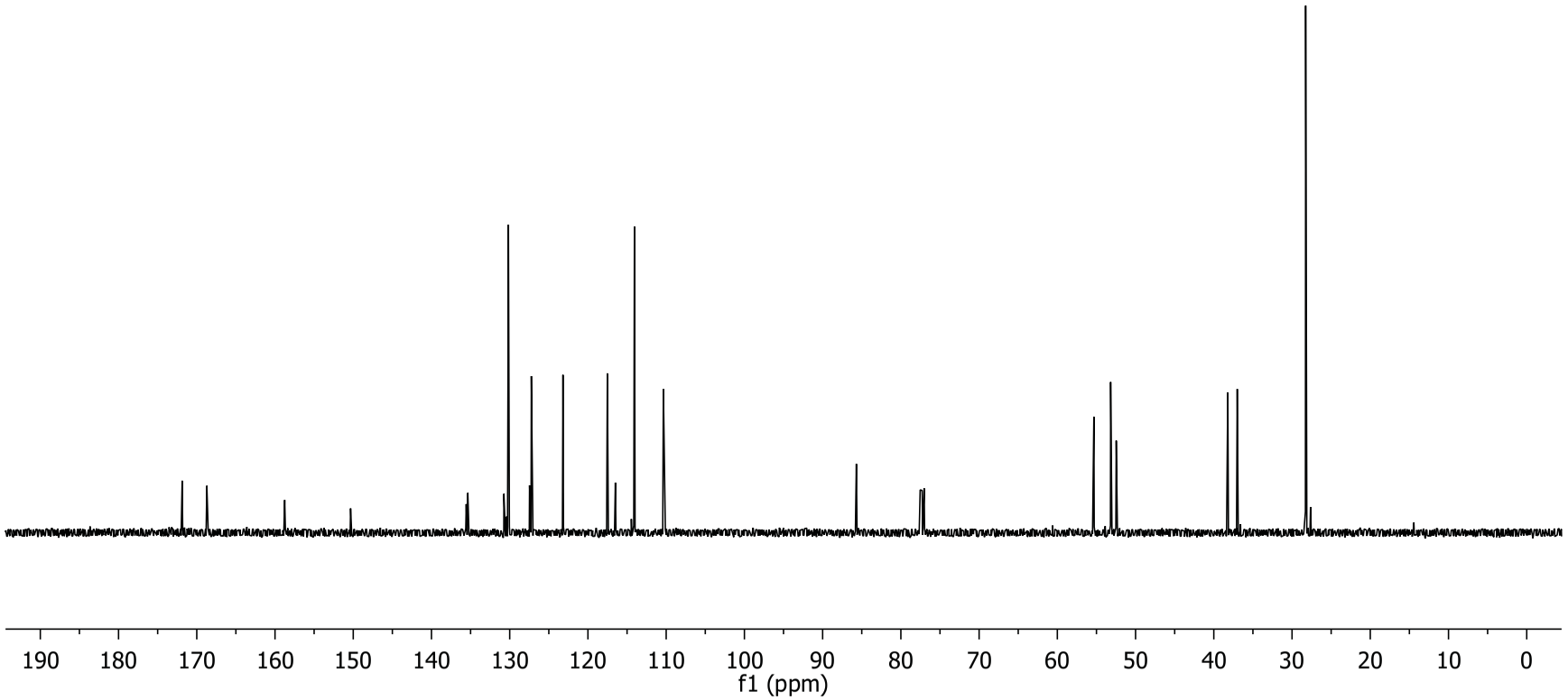

OMe
OMe
O
N
H
O
N
H
O
COOtBu
MeO
COOtBu
O
N
H
O
MeO
2.50b
0.46
1.32
1.18
1.97
3.57
0.38
0.39
0.38
1.50
0.91
3.00
6.09
3.50
1.94
3.09
2.74
1.34
1.11
21.70
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
f1 (ppm)

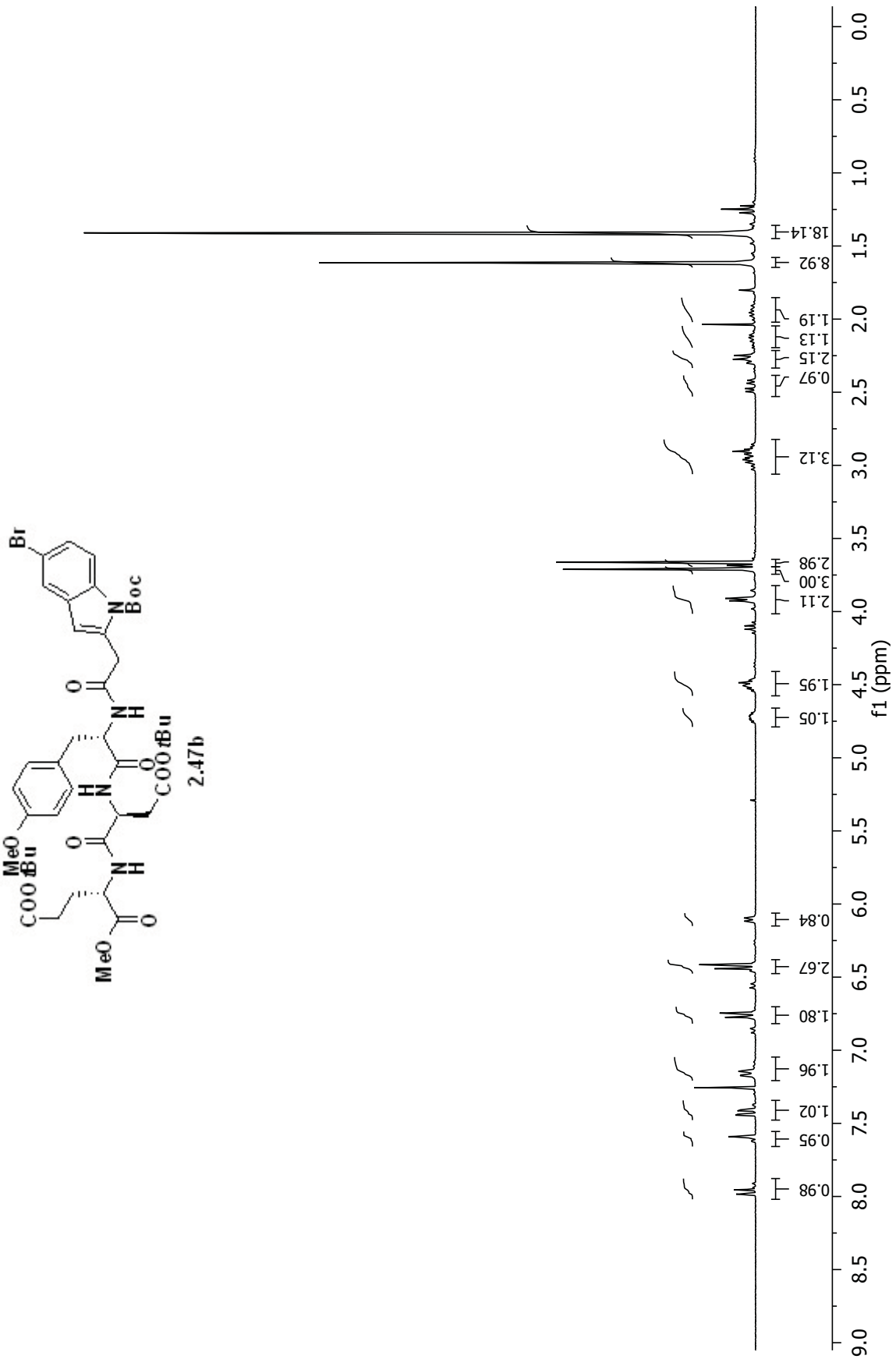

OMe
OMe
O
N
H
O
N
H
O
COOtBu
MeO
COOtBu
O
N
H
O
MeO
2.50b
f1 (ppm)

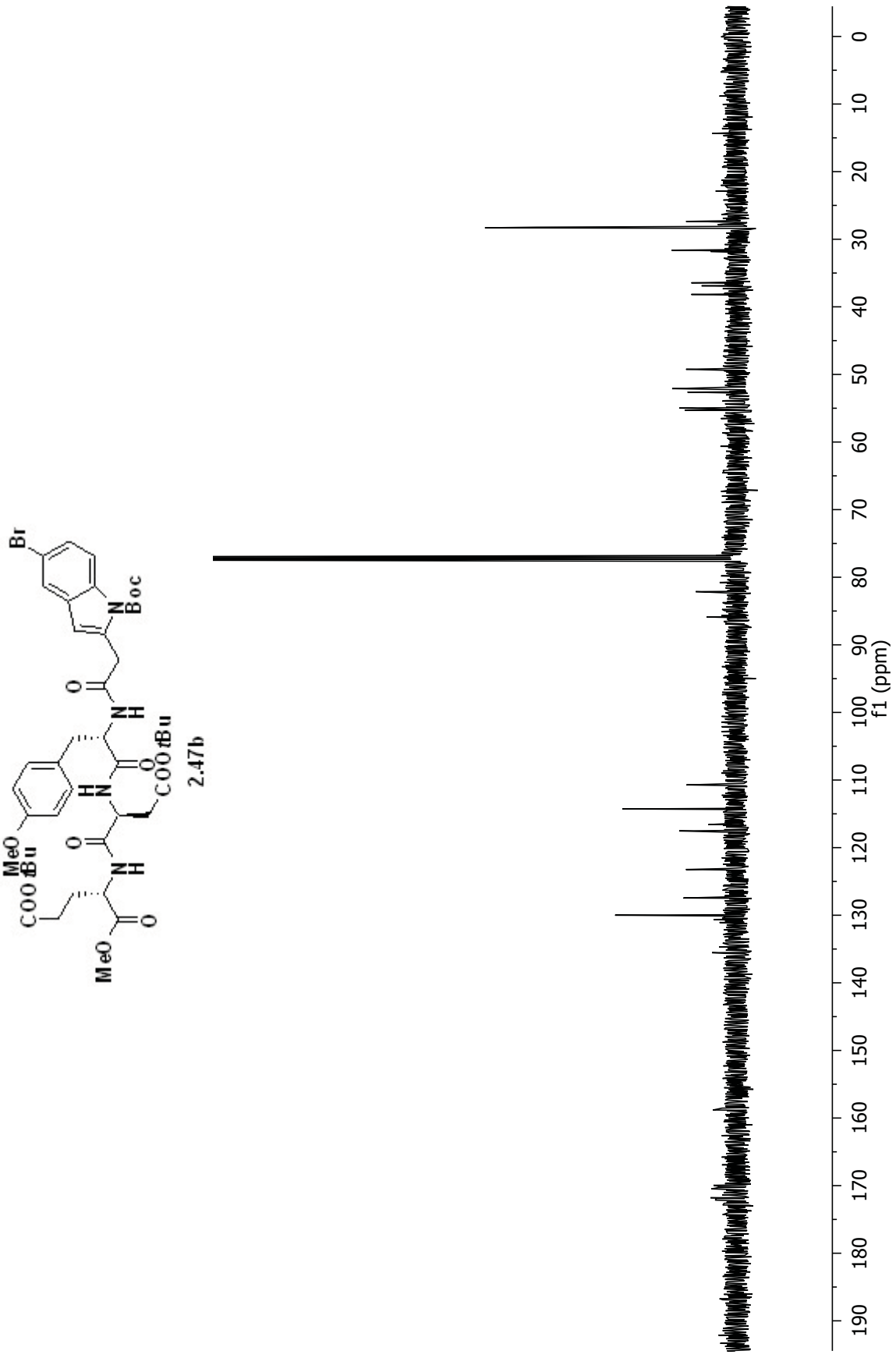

Br
N
Boc
O
N
H
MeO
O
O
MeO
2.46b
9.42
2.04
6.00
2.00
0.96
0.89
1.82
0.91
1.88
0.87
0.79
0.86
9.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
f1 (ppm)

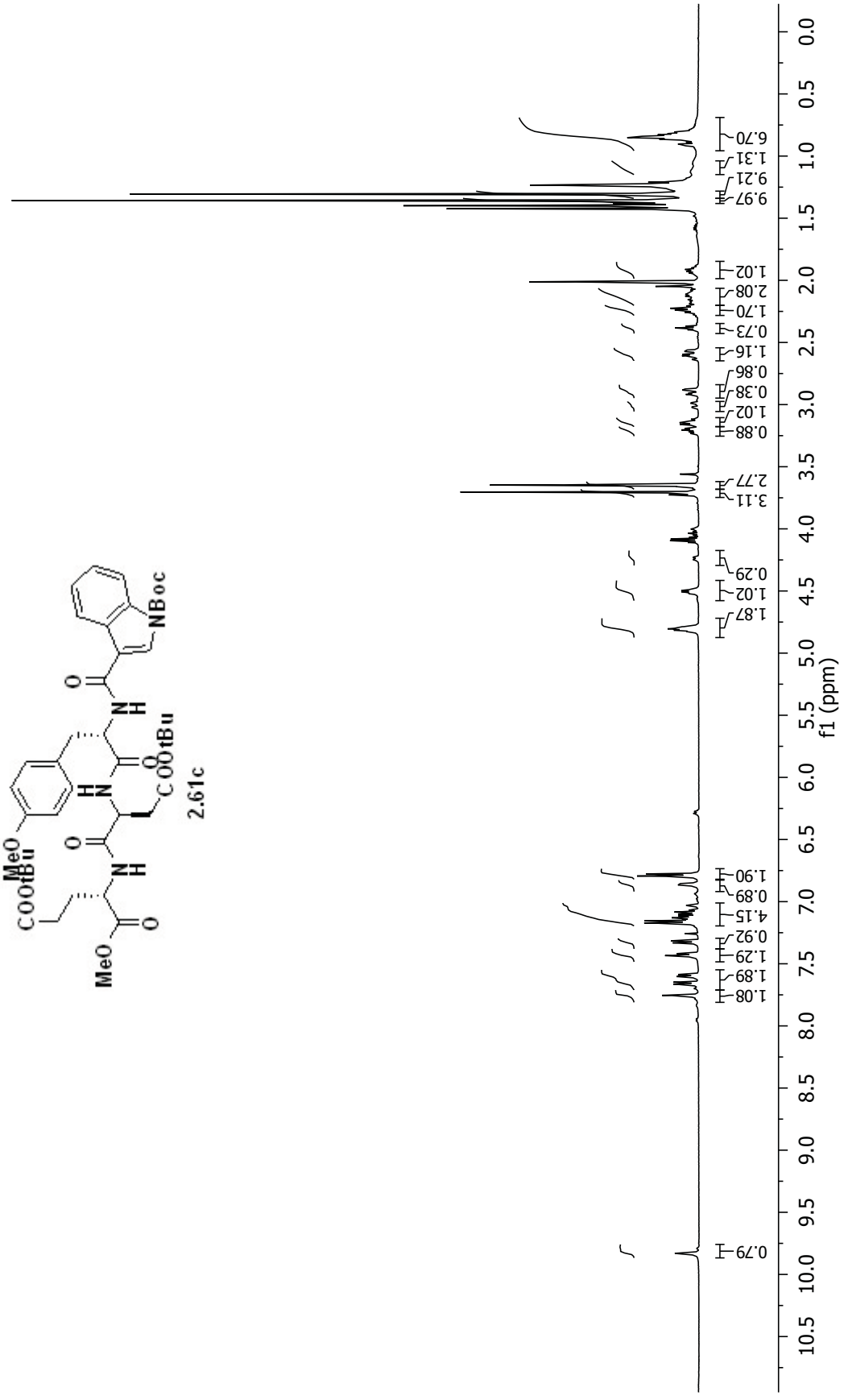

MeO
Br
O
MeO
N
H
N
Boc
O

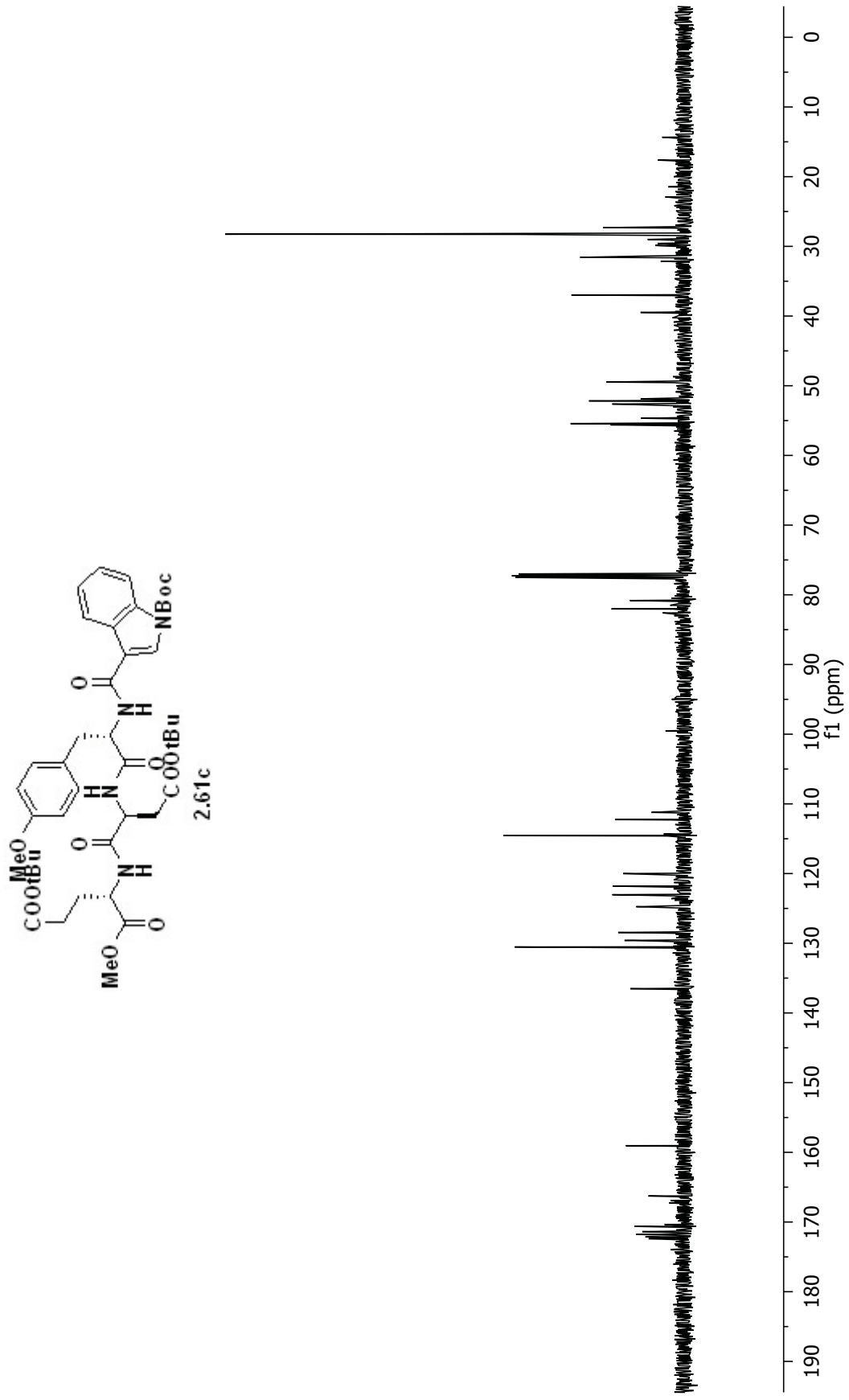
2.46b

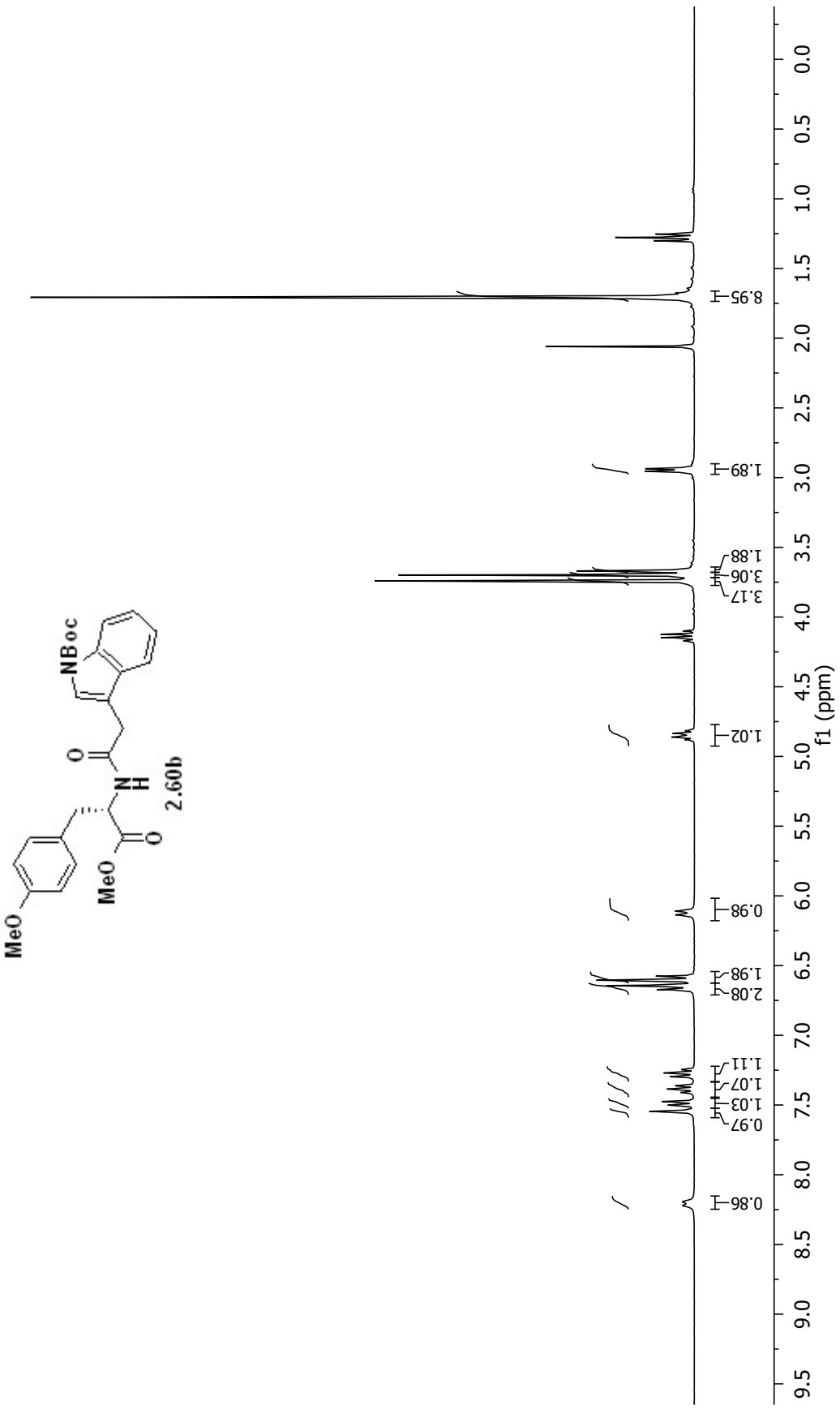

190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
f1 (ppm)

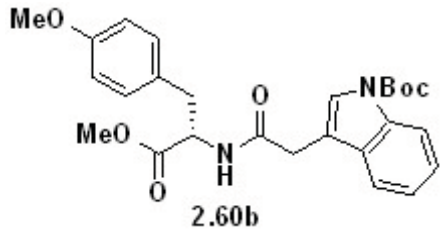

Br
N
Boc
O
N
H
O
N
H
O
O
COOtBu
MeO
COOtBu
O
N
H
O
MeO
2.47b
9.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
f1 (ppm)
0.98
0.95
1.02
1.96
1.80
2.67
0.84
1.05
1.95
2.11
3.00
2.98
3.12
0.97
2.15
1.13
1.19
8.92
18.14


Br
N
Boc
O
N
H
O
N
H
O
COOtBu
MeO
COOtBu
O
N
H
O
MeO
2.47b
f1 (ppm)
0
10
20
30
40
50
60
70
80
90
100
110
120
130
140
150
160
170
180
190

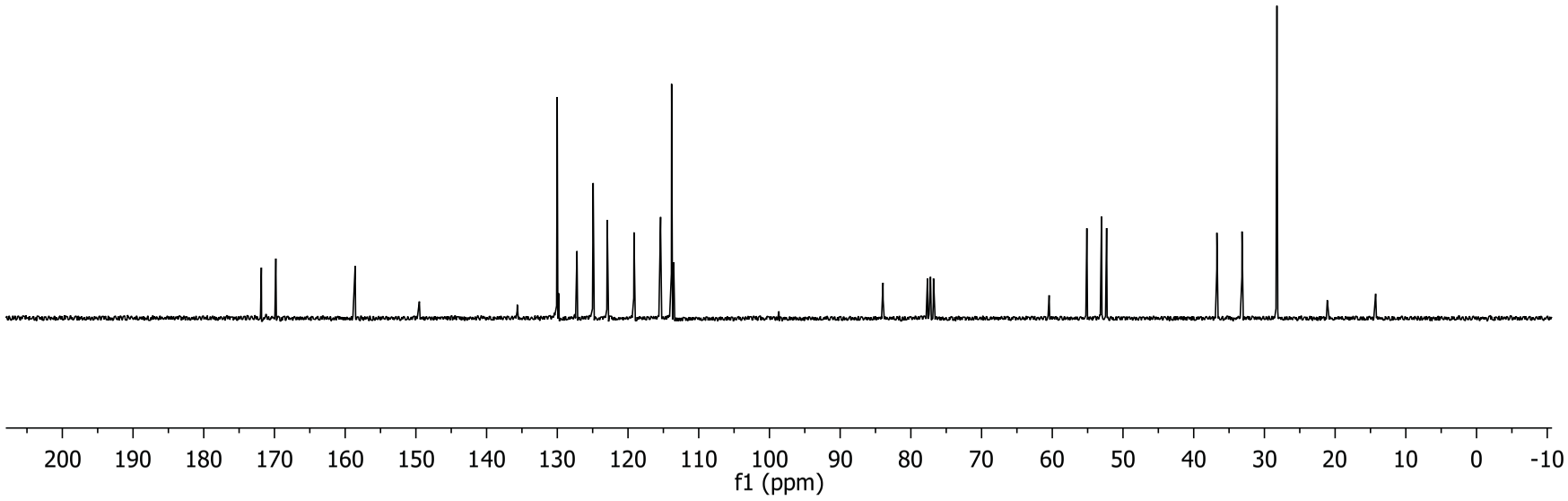

MeO
COOtBu
O
N
H
O
N
H
O
COOtBu
O
NBoc
MeO
2.61c
10.5
10.0
9.5
9.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
f1 (ppm)
0.79
1.80
1.89
1.29
0.92
4.15
0.89
1.90
1.87
1.02
0.29
3.11
2.77
0.88
1.02
0.38
0.86
1.16
0.73
1.70
2.08
1.02
9.97
9.21
1.31
6.70

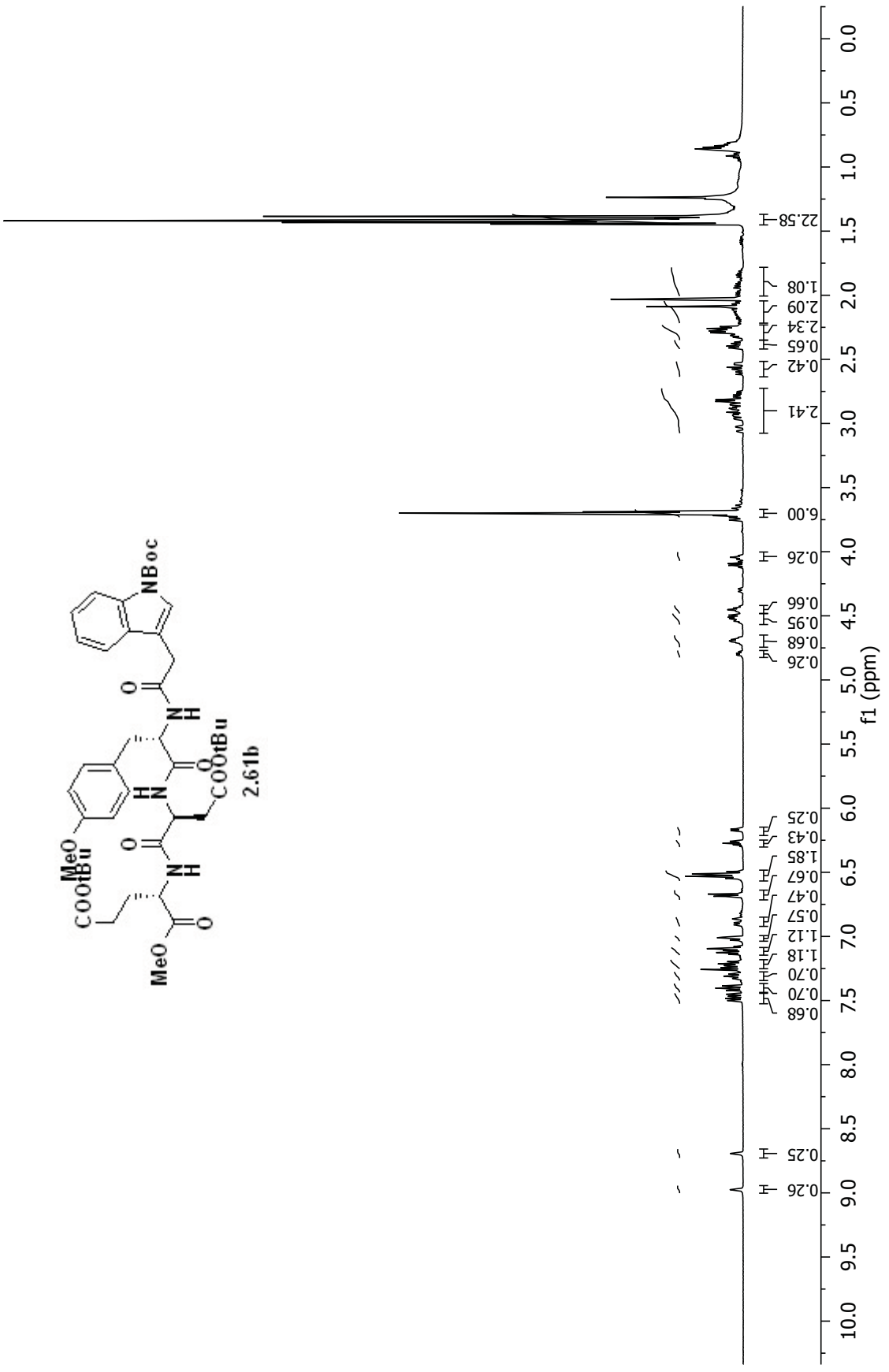

MeO
COOtBu
O
N
H
O
N
H
O
COOtBu
O
NBoc
MeO
O
2.61c
f1 (ppm)

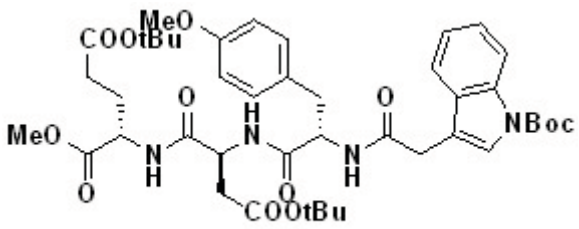

MeO
MeO
O
N
H
O
NBoc
2.60b
8.95
1.89
1.88
3.06
3.17
1.02
0.98
2.08
1.98
1.11
1.07
1.03
0.97
0.86
f1 (ppm)


MeO
MeO
O
N
H
O
NBoc

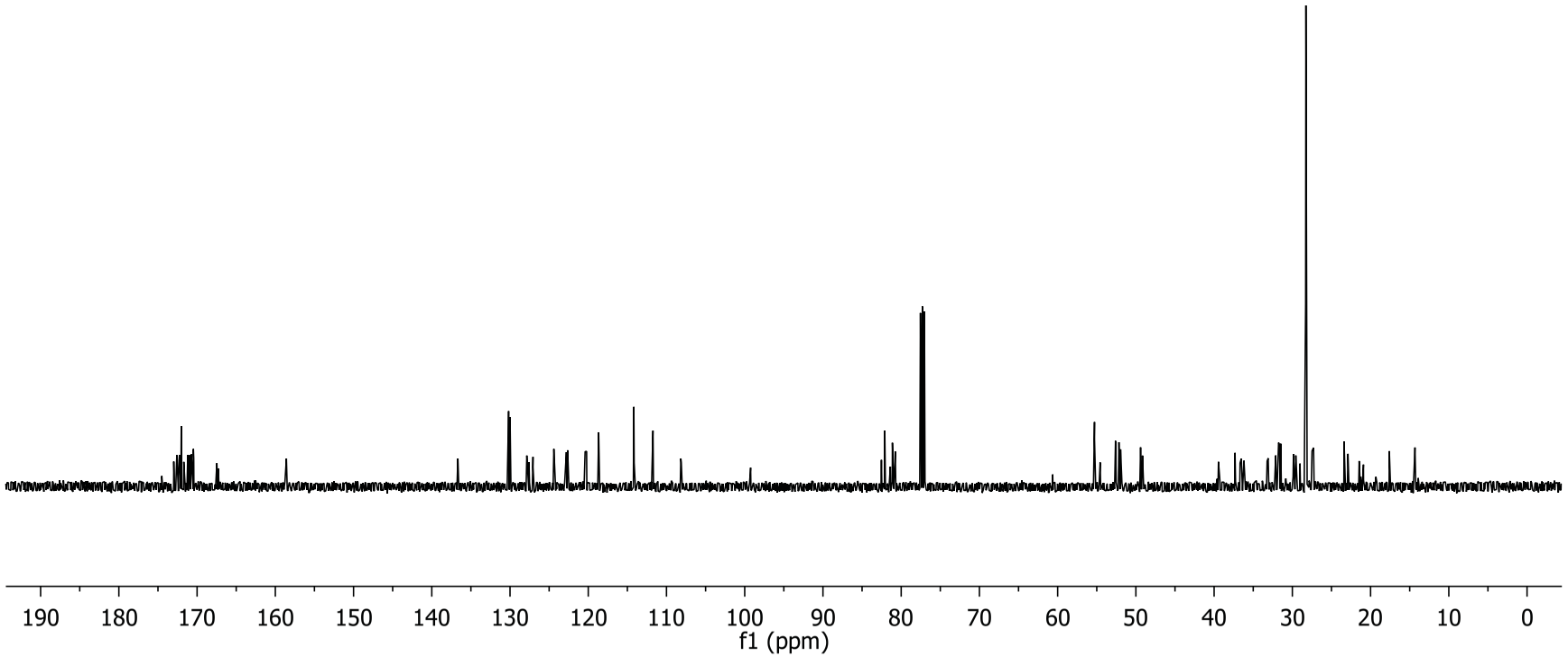
2.60b

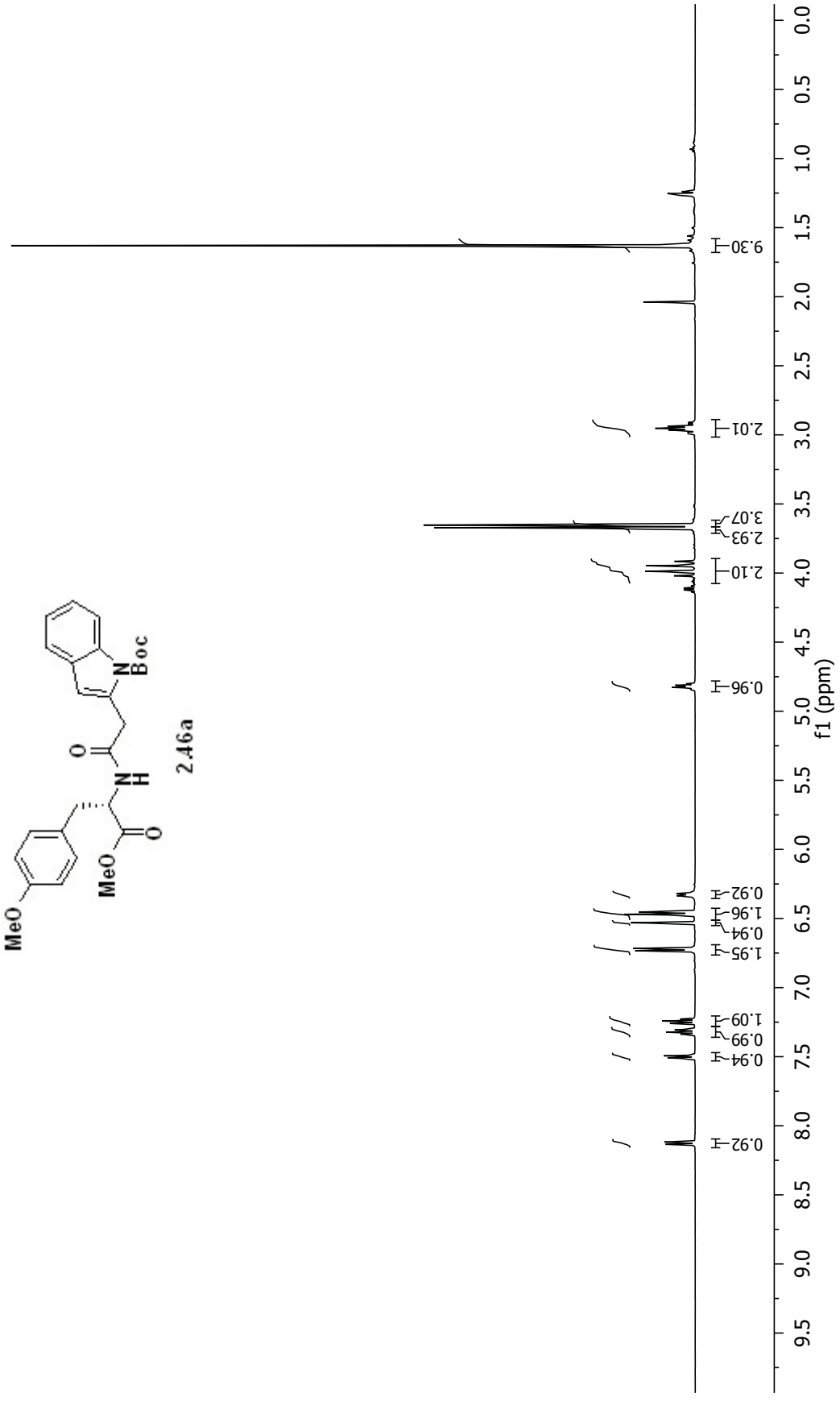

200
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
-10
f1 (ppm)

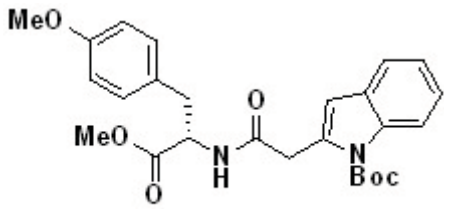

NBoc
O
N
H
O
N
H
O
COOtBu
MeO
COOtBu
O
N
H
O
MeO
2.61b
0.26
0.25
0.68
0.70
0.70
1.18
1.12
0.57
0.47
0.67
1.85
0.43
0.25
0.26
0.68
0.95
0.66
0.26
6.00
0.42
2.41
0.65
2.34
2.09
1.08
22.58
10.0
9.5
9.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
f1 (ppm)


MeO
COOtBu
MeO
O
H
N
O
NBoc
N
H
O
N
H
O
COOtBu

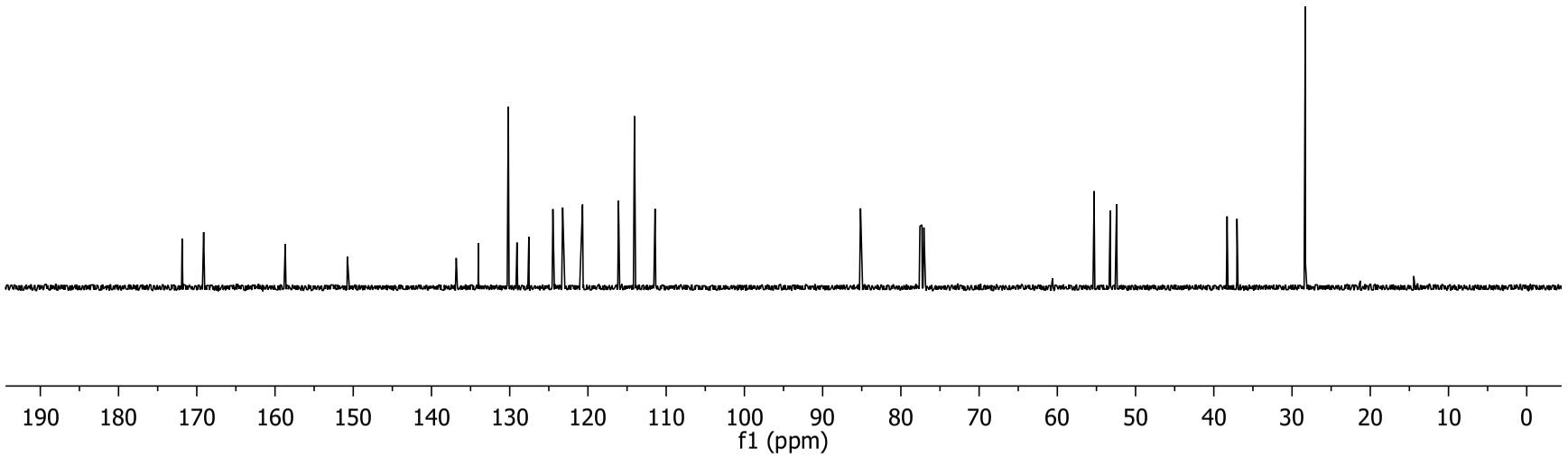
2.61b

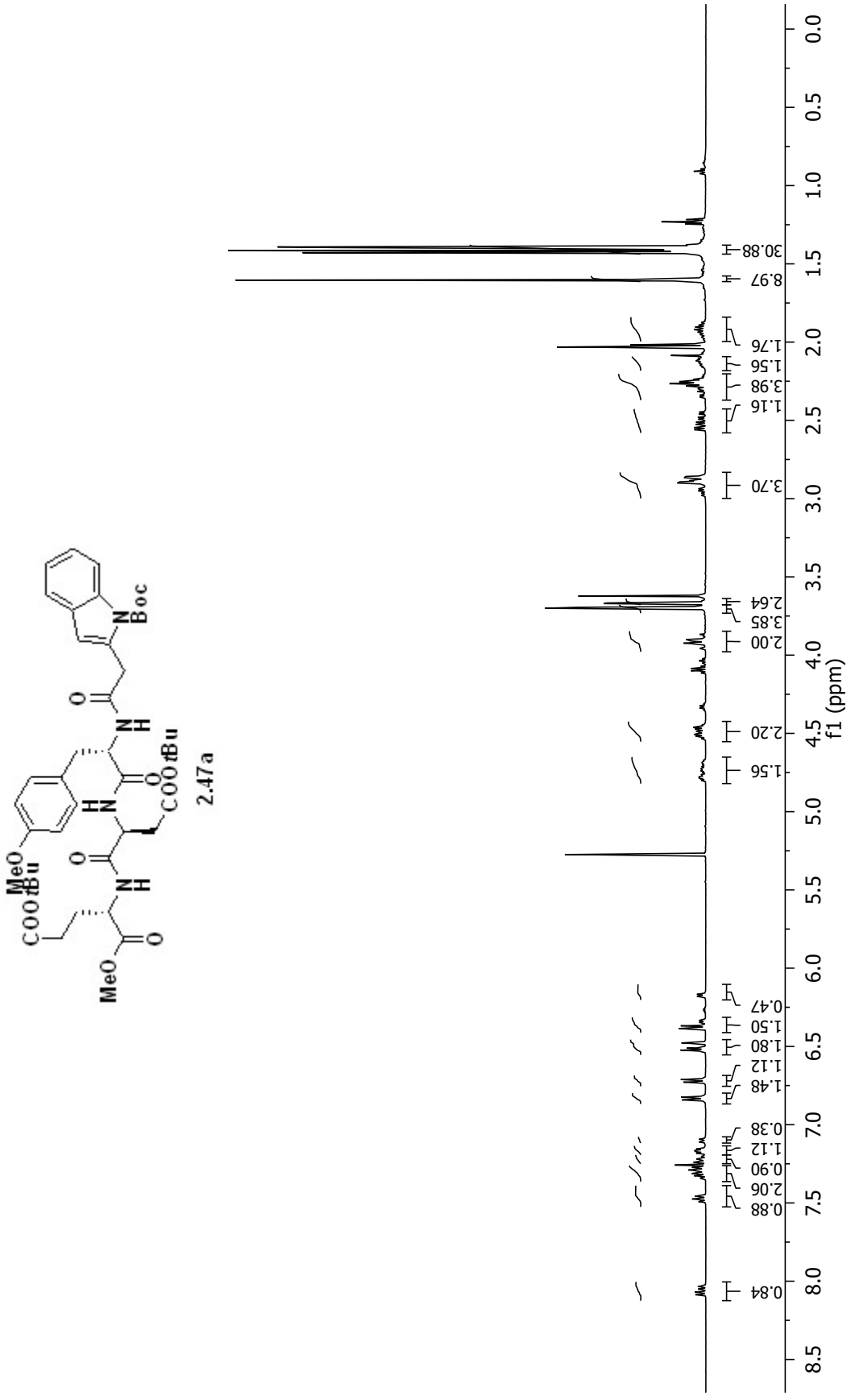

190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
f1 (ppm)

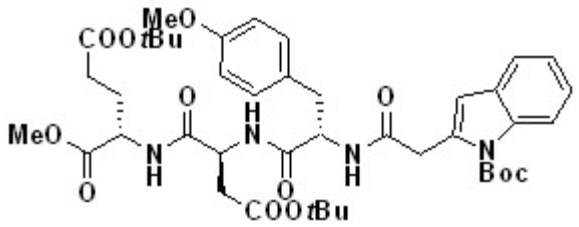

MeO
MeO
O
O
N
H
O
N
Boc
2.46a
9.30
2.01
3.07
2.93
2.10
0.96
0.92
1.96
0.94
1.95
1.09
0.99
0.94
0.92
9.5
9.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
f1 (ppm)


MeO
MeO
O
O
N
H
N
Boc

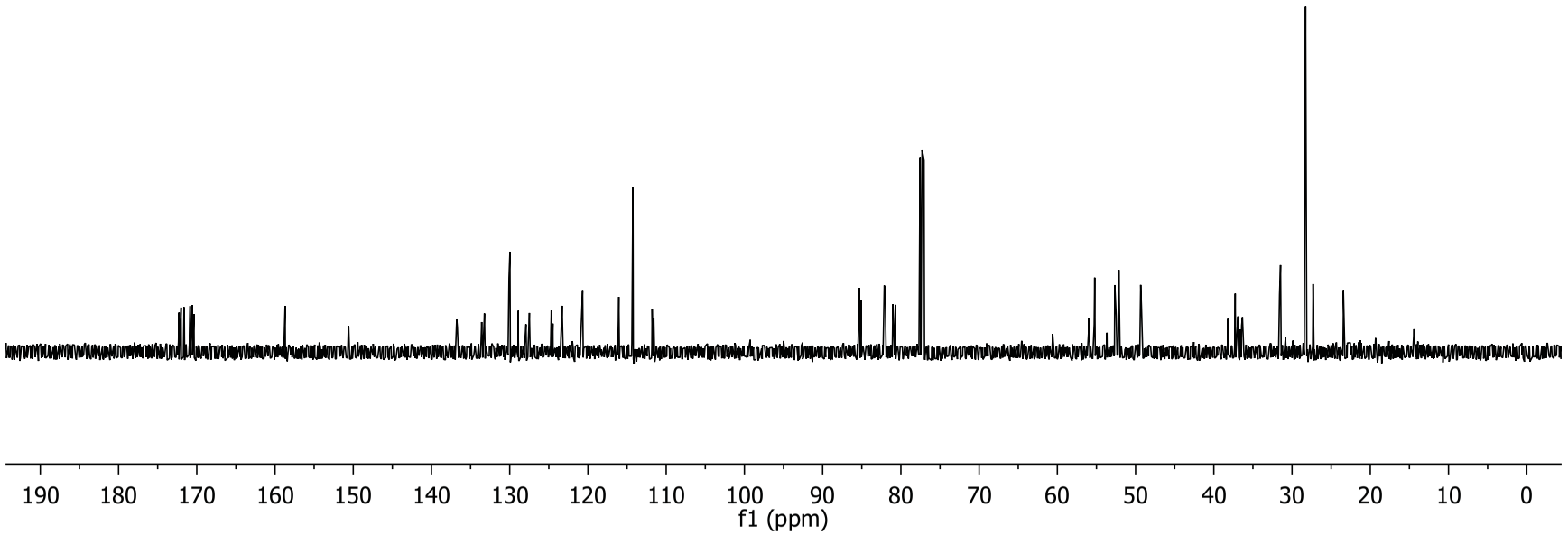
2.46a

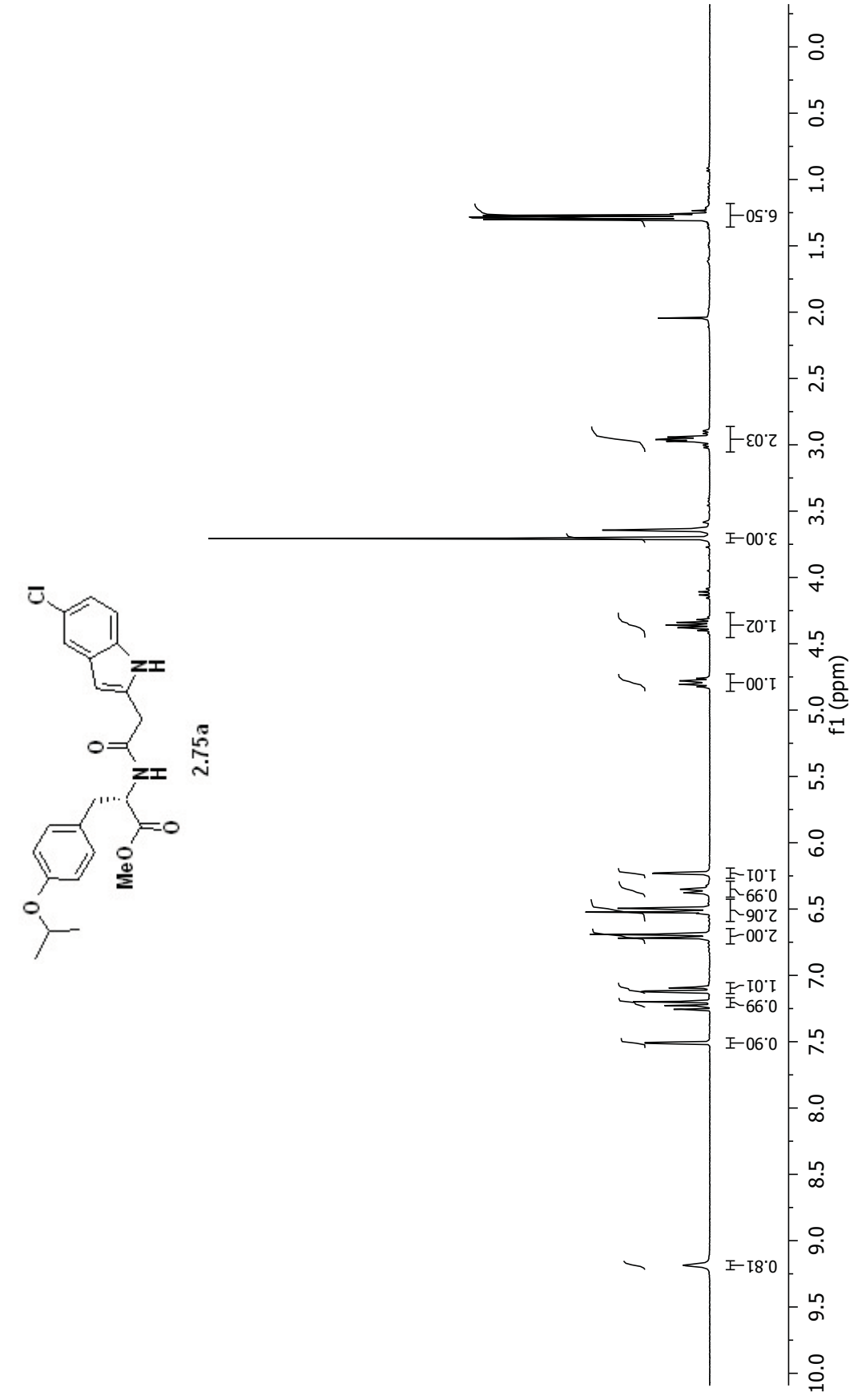

190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
f1 (ppm)

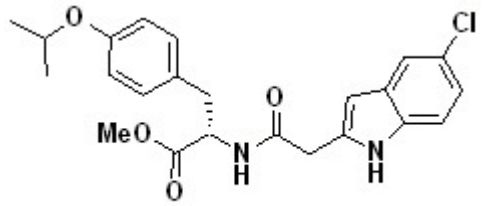

MeO
COOtBu
O
N
H
O
N
H
O
COOtBu
O
N
Boc
MeO
COOtBu
2.47a
0.84
0.88
2.06
0.90
1.12
0.38
1.48
1.12
1.80
1.50
0.47
1.56
2.20
2.00
3.85
2.64
3.70
1.16
3.98
1.56
1.76
8.97
30.88
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
f1 (ppm)


MeO
COOtBu
MeO
O
H
N
O
N
H
O
N
H
O
N
Boc
O
COOtBu

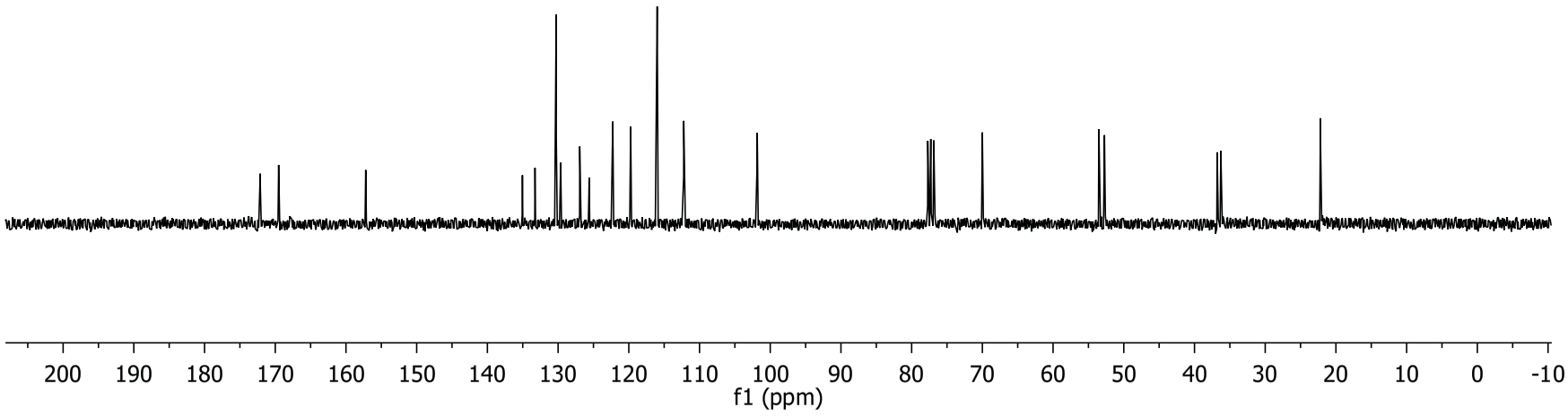
2.47a

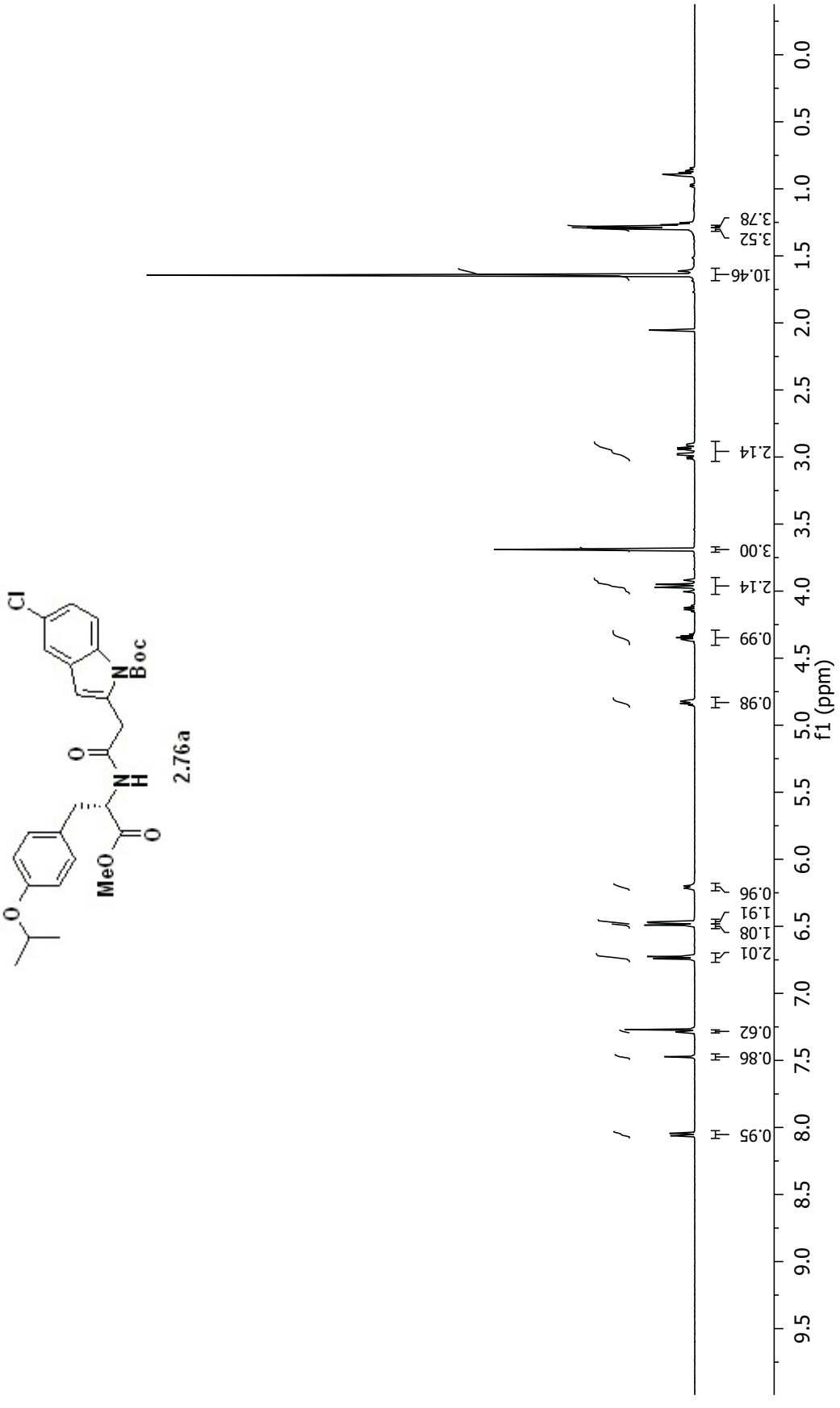

190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
f1 (ppm)

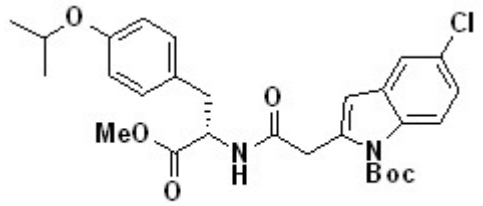

Cl
N
H
O
N
H
MeO
O
O
2.75a
6.50
2.03
3.00
1.02
1.00
1.01
0.99
2.06
2.00
1.01
0.99
0.90
0.81
10.0
9.5
9.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
f1 (ppm)


O
Cl
MeO
O
N
H
O
N
H
O

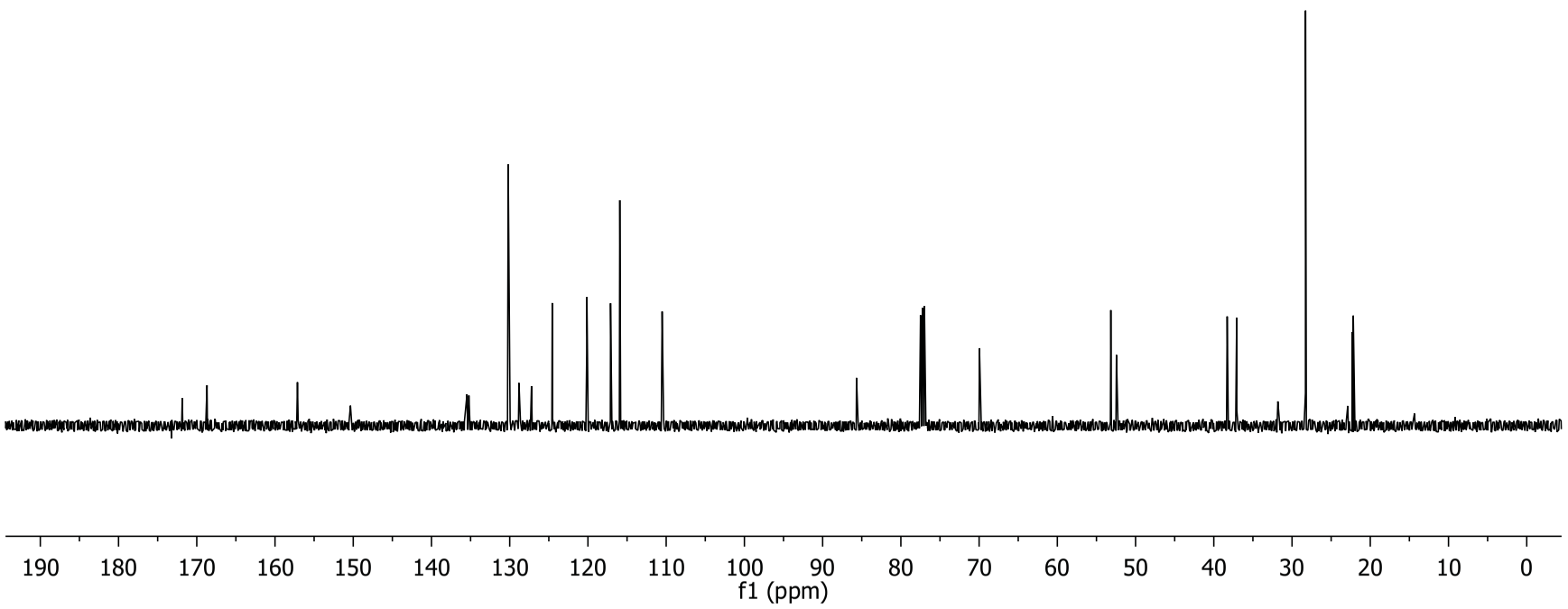
2.75a

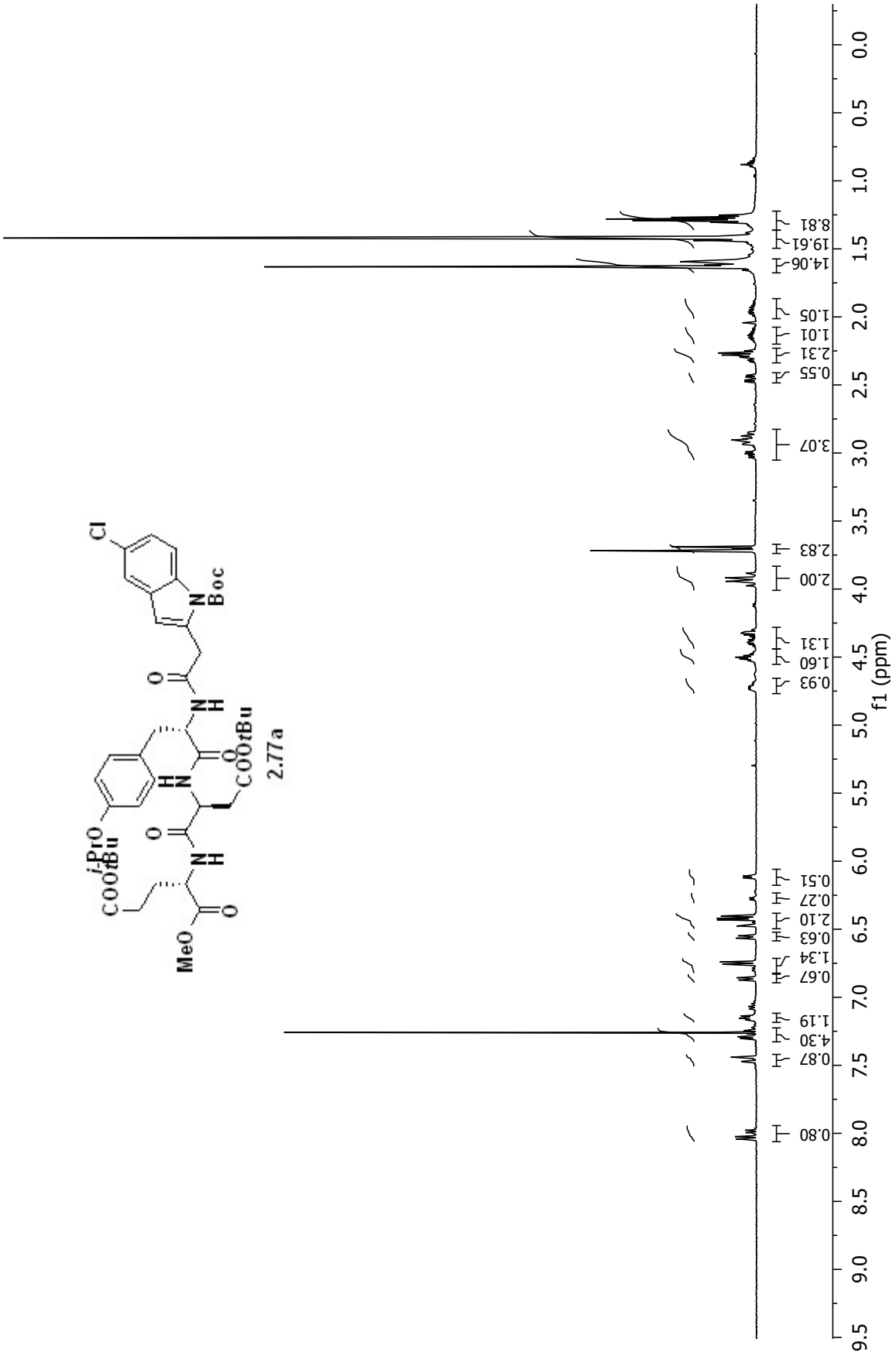

200
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
-10
f1 (ppm)

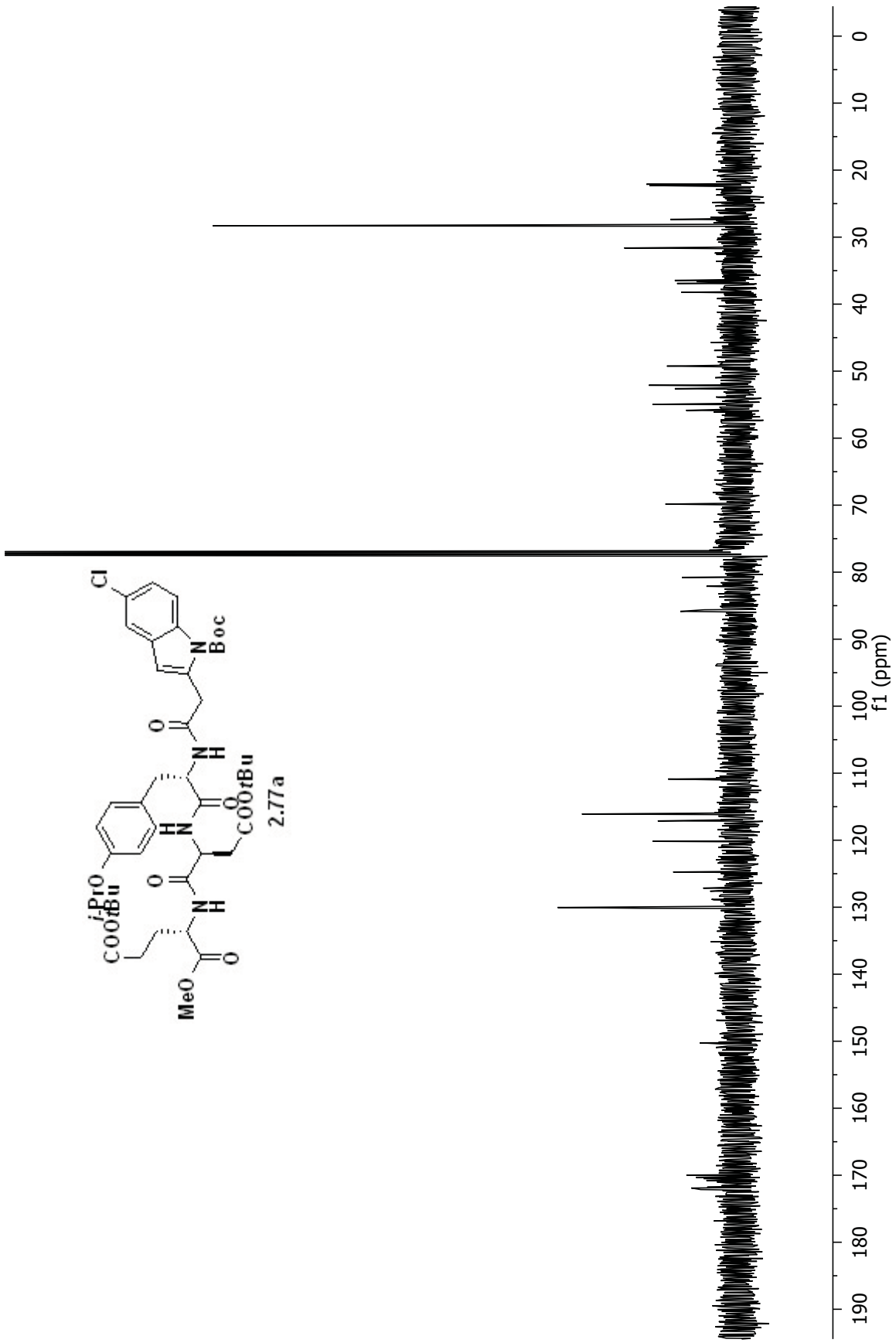

Cl
N
Boc
O
N
H
MeO
O
O
2.76a
0.95
0.86
0.62
2.01
1.08
1.91
0.96
0.98
0.99
2.14
3.00
2.14
10.46
3.52
3.78
9.5
9.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
f1 (ppm)

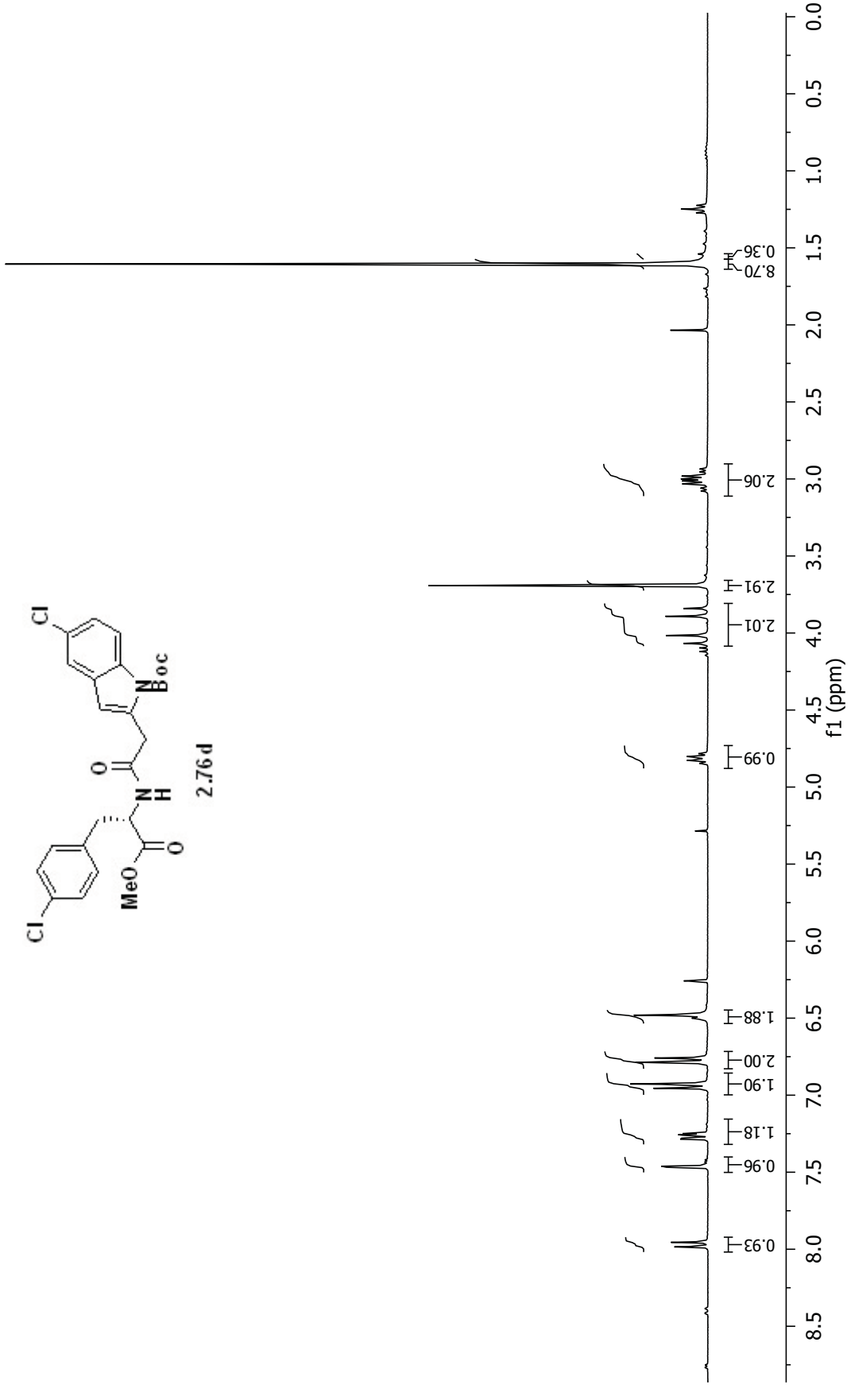

O
Cl
MeO
O
N
H
O
N
Boc
O

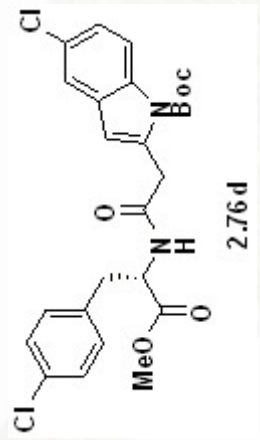
2.76a


190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
f1 (ppm)

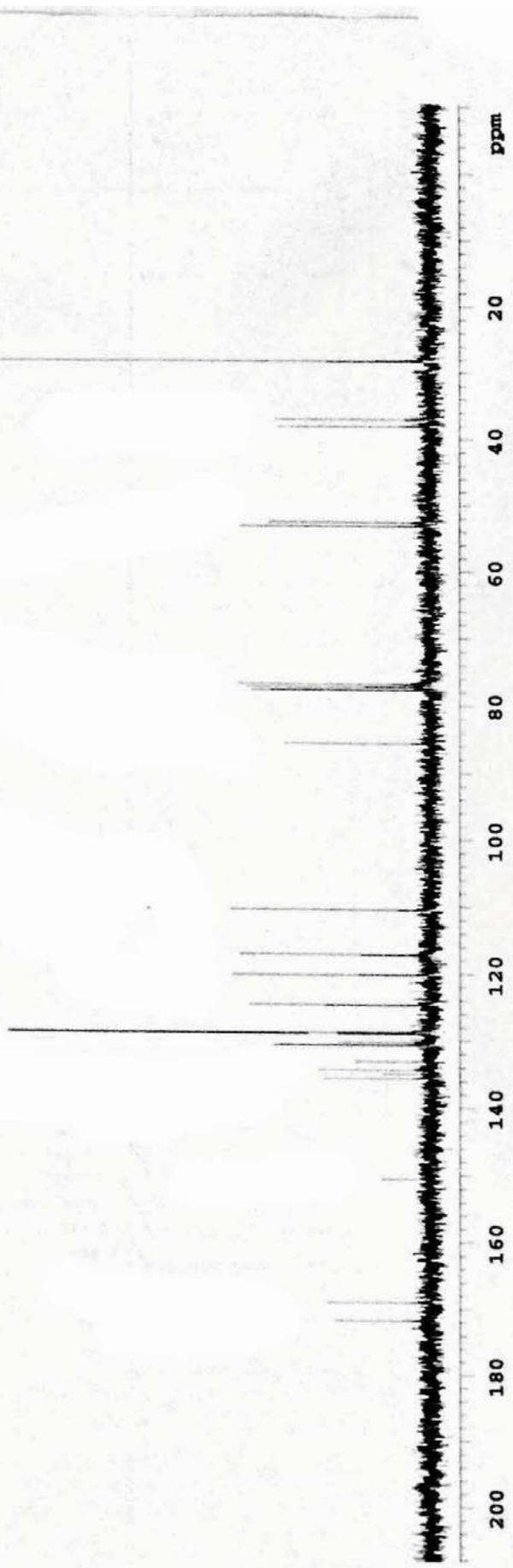

Cl
N
Boc
O
N
H
O
N
H
O
COOtBu
O
iPrO
COOtBu
O
N
H
O
MeO
2.77a
0.80
0.87
4.30
1.19
0.67
1.34
0.63
2.10
0.27
0.51
3.07
0.55
2.31
1.01
1.05
14.06
19.61
8.81
2.83
2.00
1.31
1.60
0.93
9.5
9.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
f1 (ppm)

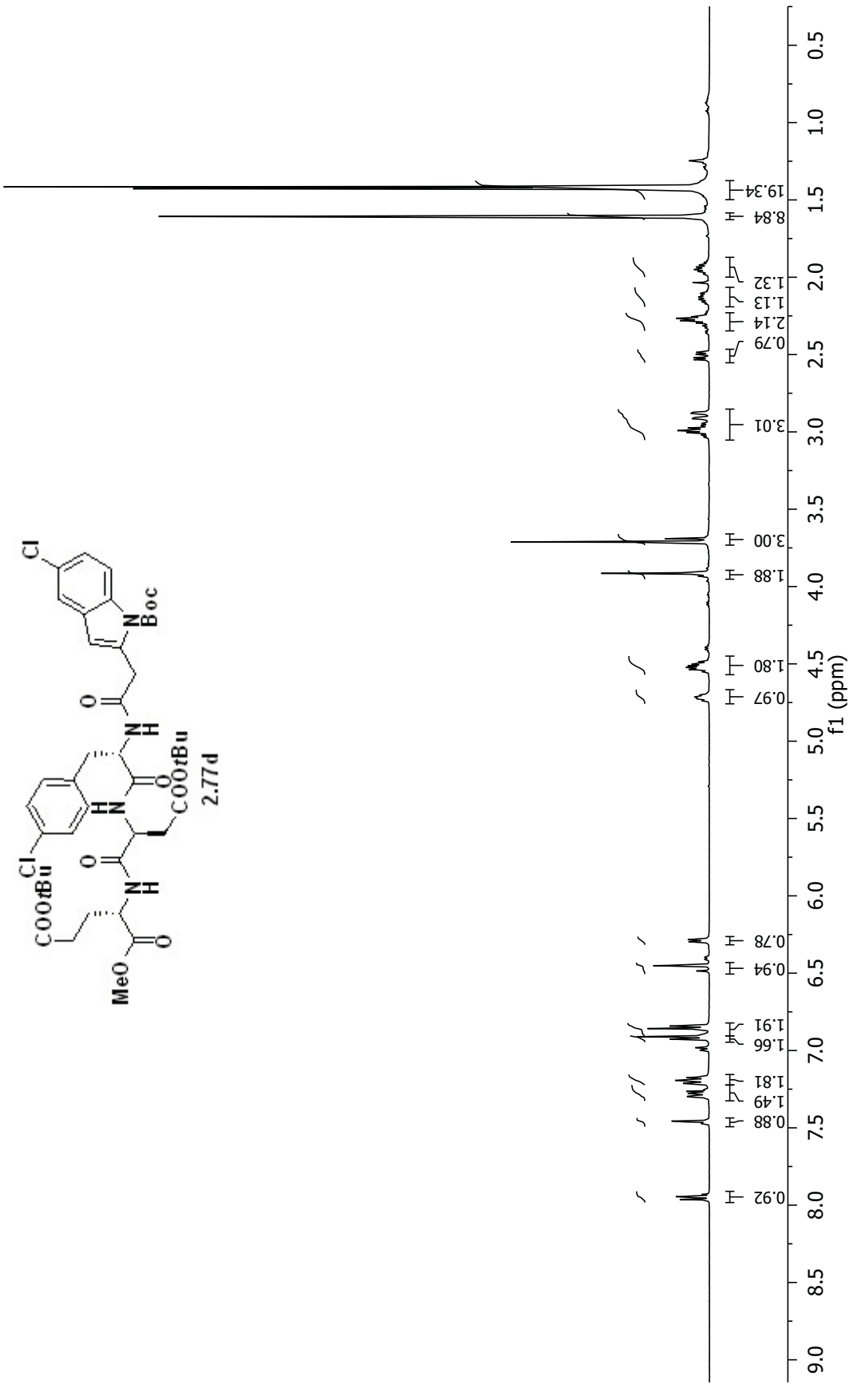

Cl
Boc
N
O
H
N
O
O
tBu
O
2.77a
H
N
O
iPrO
tBu
O
O
MeO
f1 (ppm)
0
10
20
30
40
50
60
70
80
90
100
110
120
130
140
150
160
170
180
190

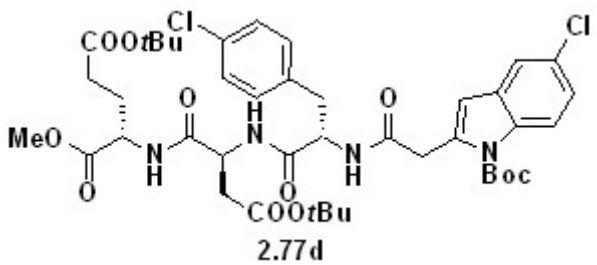

Cl
N
Boc
O
N
H
MeO
O
O
Cl
2.76d
8.70
0.36
2.06
2.91
2.01
0.99
1.88
2.00
1.90
1.18
0.96
0.93
f1 (ppm)



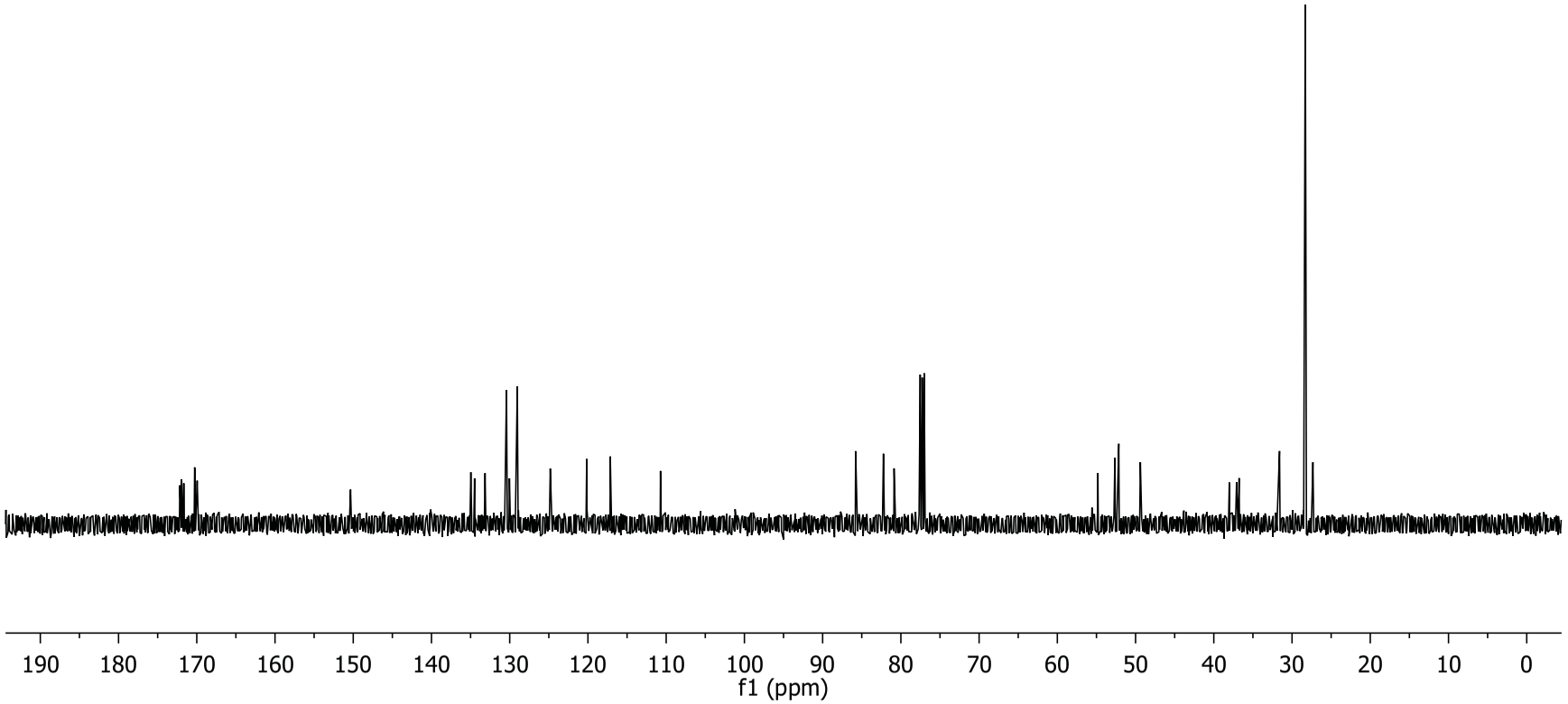
2.76d

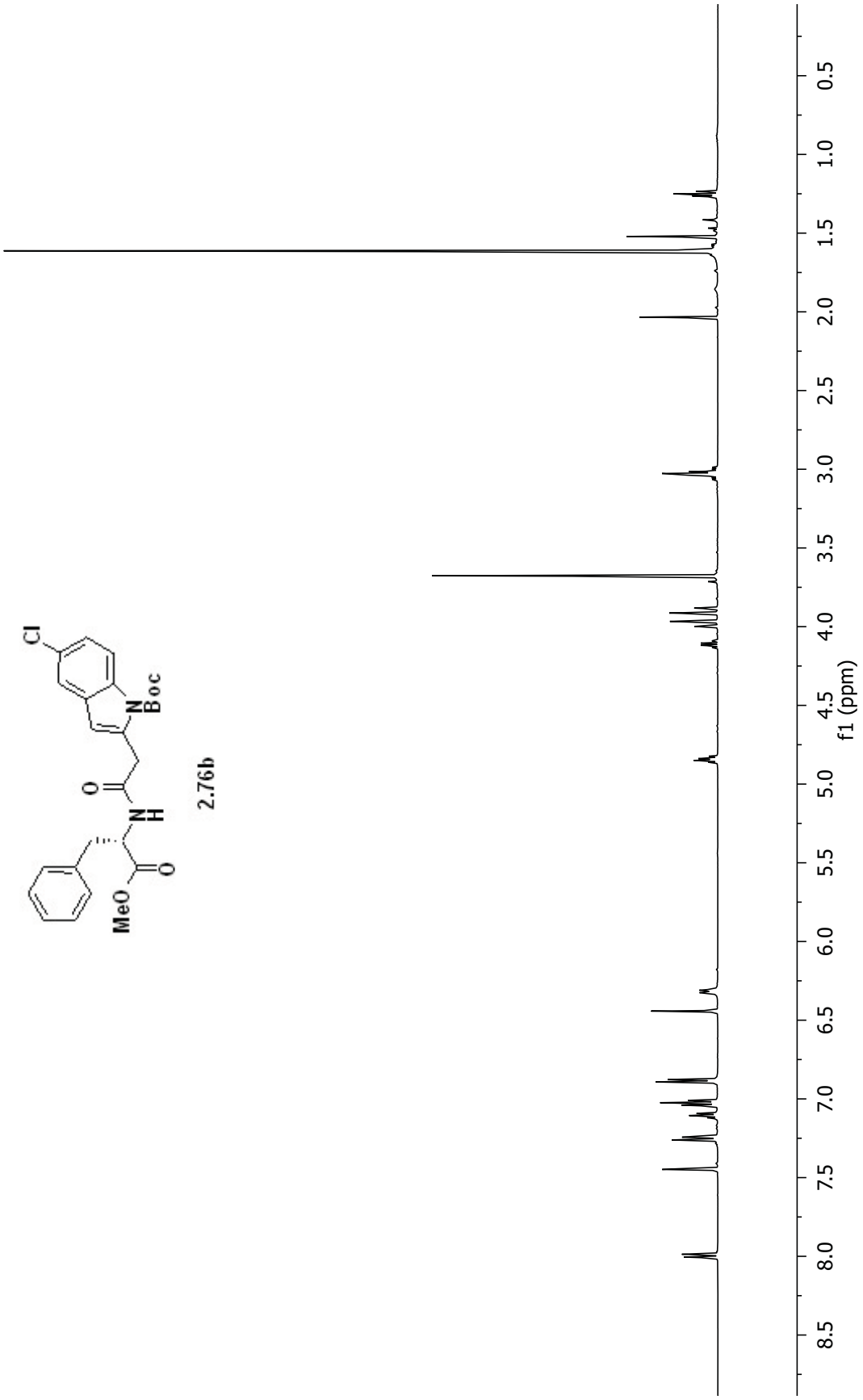


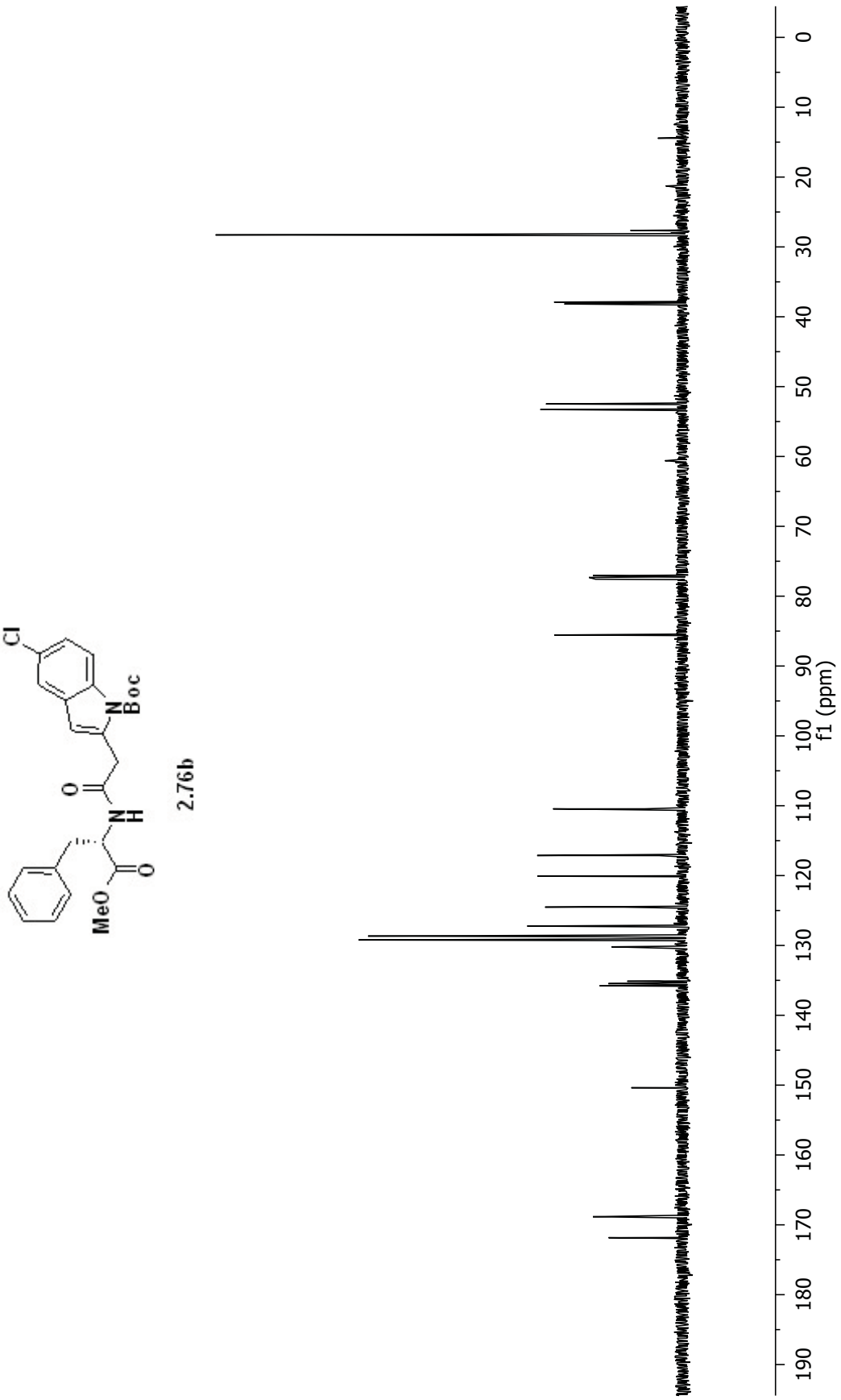

Cl
N
Boc
O
N
H
O
N
H
O
COOtBu
2.77 d
Cl
COOtBu
O
N
H
O
MeO
9.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
f1 (ppm)
0.92
0.88
1.49
1.81
1.66
1.91
0.94
0.78
0.97
1.80
1.88
3.00
3.01
0.79
2.14
1.13
1.32
8.84
19.34

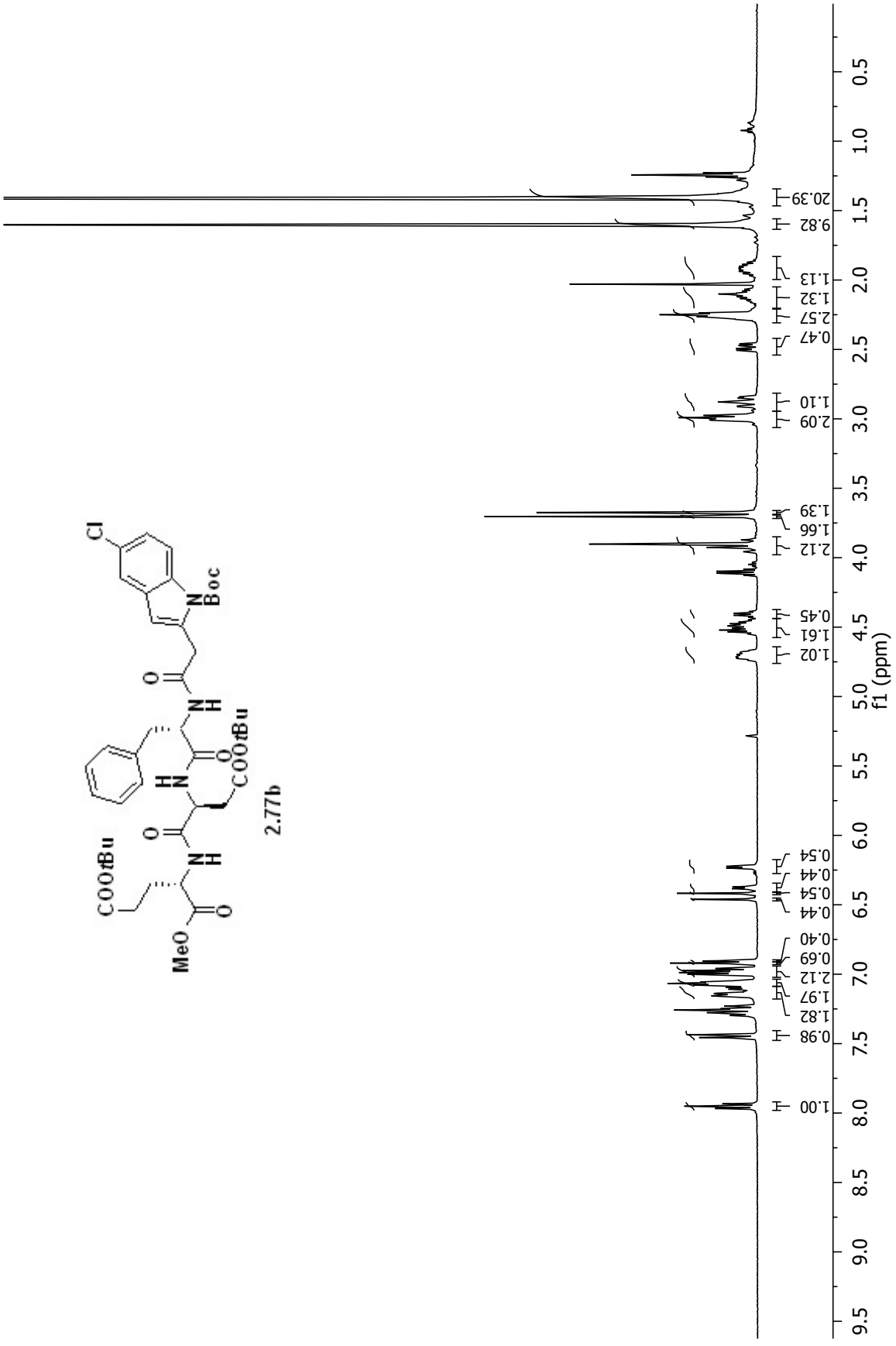

Cl
COOtBu
MeO
O
NH
O
H
N
O
NH
O
Cl
N
Boc
COOtBu

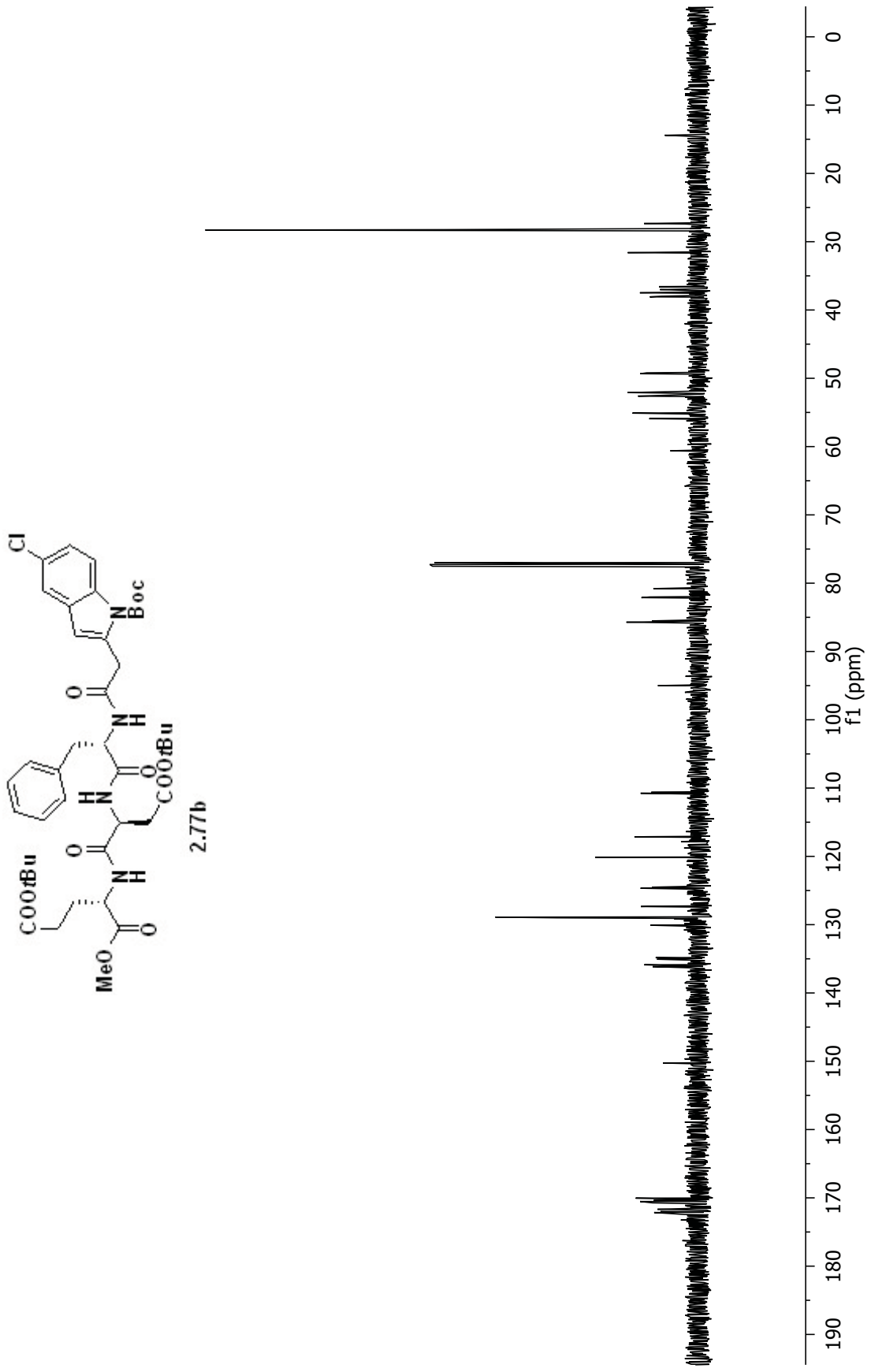
2.77d

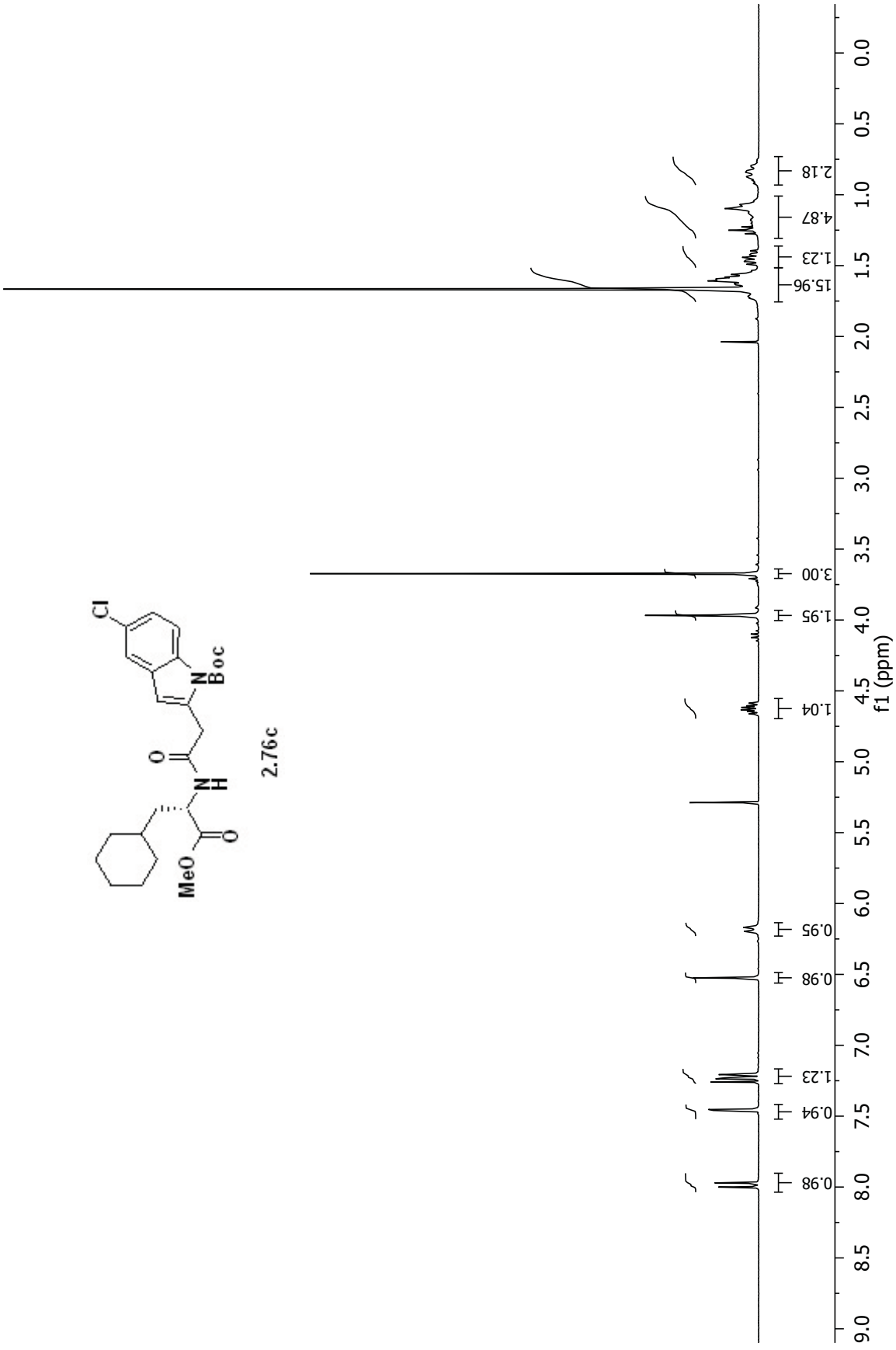

190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
f1 (ppm)

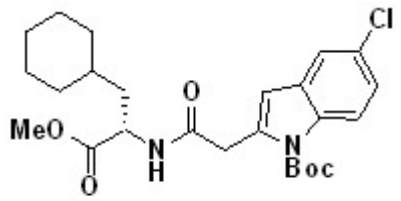

Cl
N
Boc
O
N
H
MeO
O
2.76b
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
f1 (ppm)


Cl
N
Boc
O
N
H
O
MeO
O
2.76b
f1 (ppm)
0
10
20
30
40
50
60
70
80
90
100
110
120
130
140
150
160
170
180
190

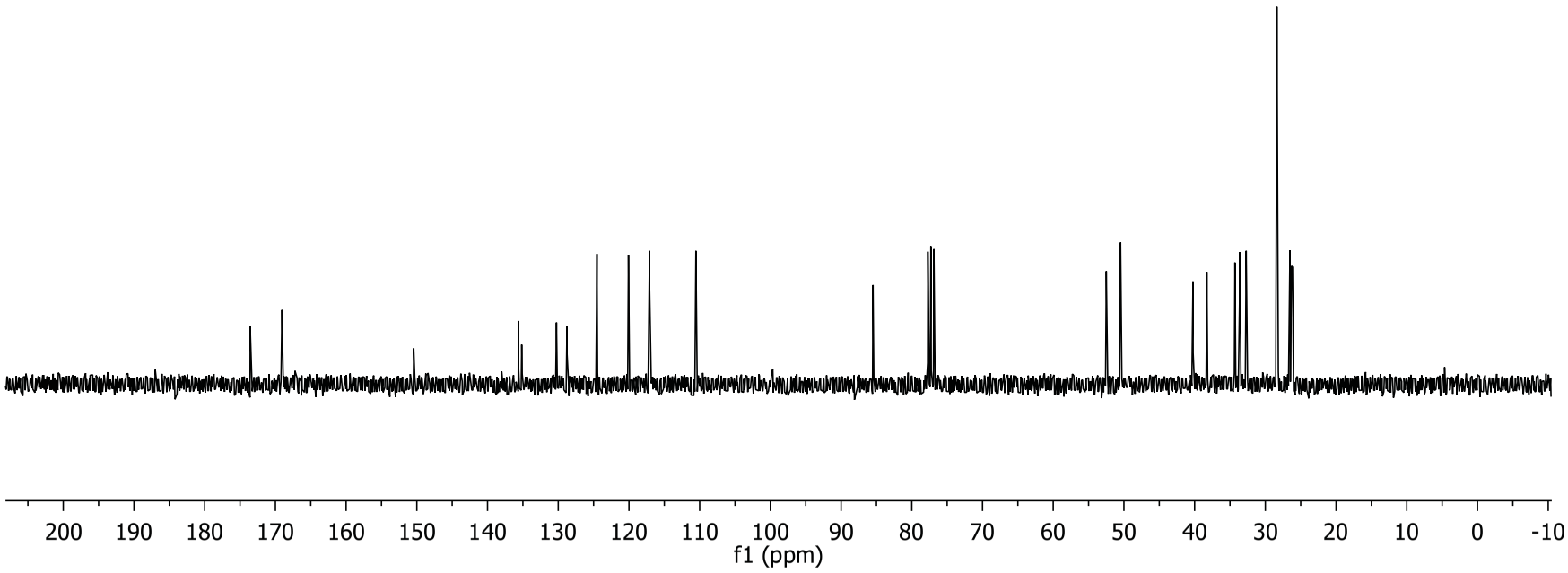

Cl
N
Boc
O
N
H
O
N
H
O
COOtBu
COOtBu
O
N
H
O
MeO
O
2.77b
1.00
0.98
1.82
1.97
2.12
0.69
0.40
0.44
0.54
0.44
0.54
1.02
1.61
0.45
2.12
1.66
1.39
2.09
1.10
0.47
2.57
1.32
1.13
9.82
20.39
9.5
9.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
f1 (ppm)


Cl
N
Boc
O
N
H
O
N
H
O
COOtBu
COOtBu
O
N
H
O
MeO
2.77b
0
10
20
30
40
50
60
70
80
90
100
110
120
130
140
150
160
170
180
190
f1 (ppm)

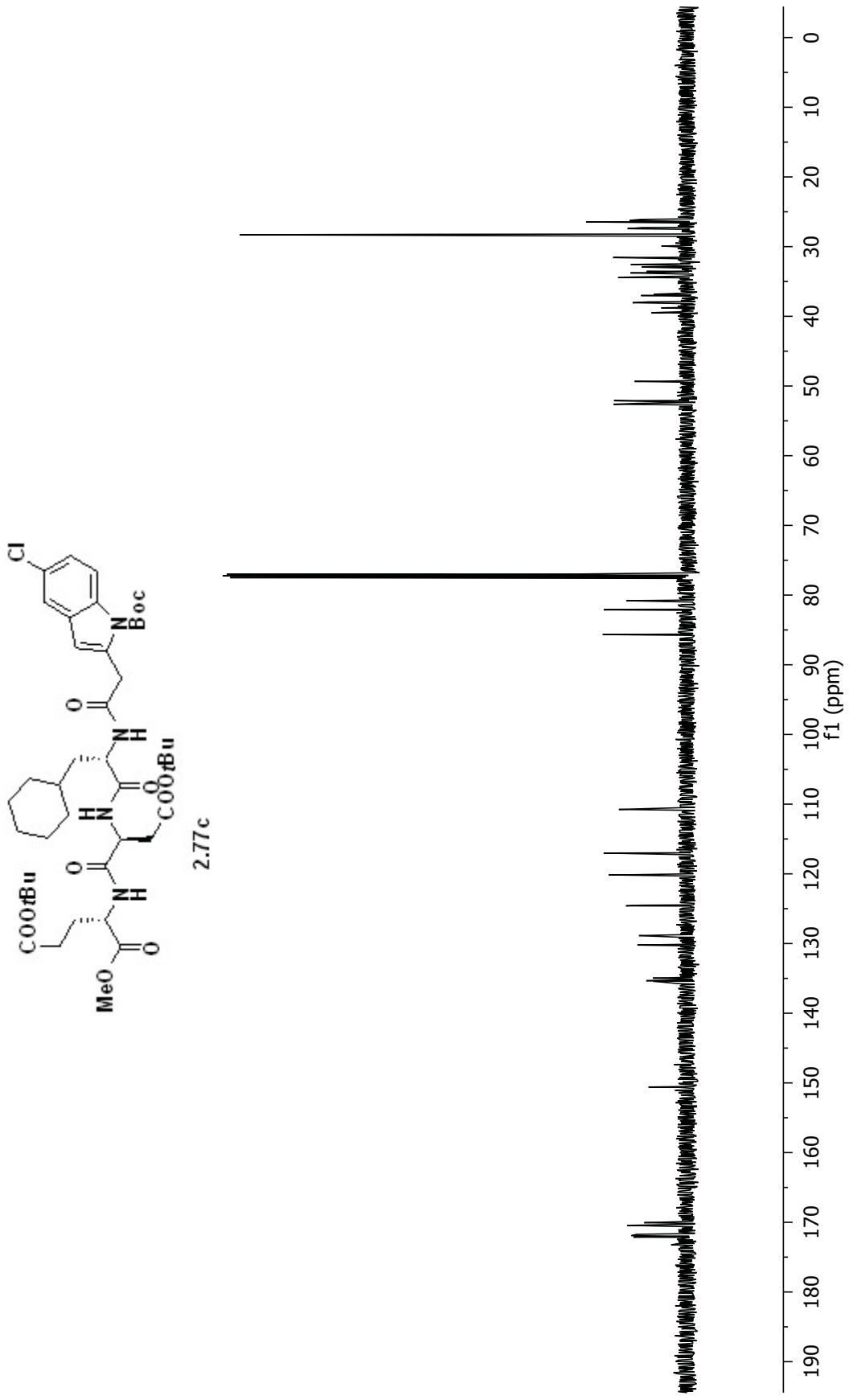

Cl
N
Boc
O
N
H
O
MeO
O
2.76c
0.98
0.94
1.23
0.98
0.95
6.0
5.5
5.0
1.04
4.5
4.0
1.95
3.00
3.5
3.0
2.5
2.0
15.96
1.5
1.23
4.87
1.0
2.18
0.5
0.0
f1 (ppm)

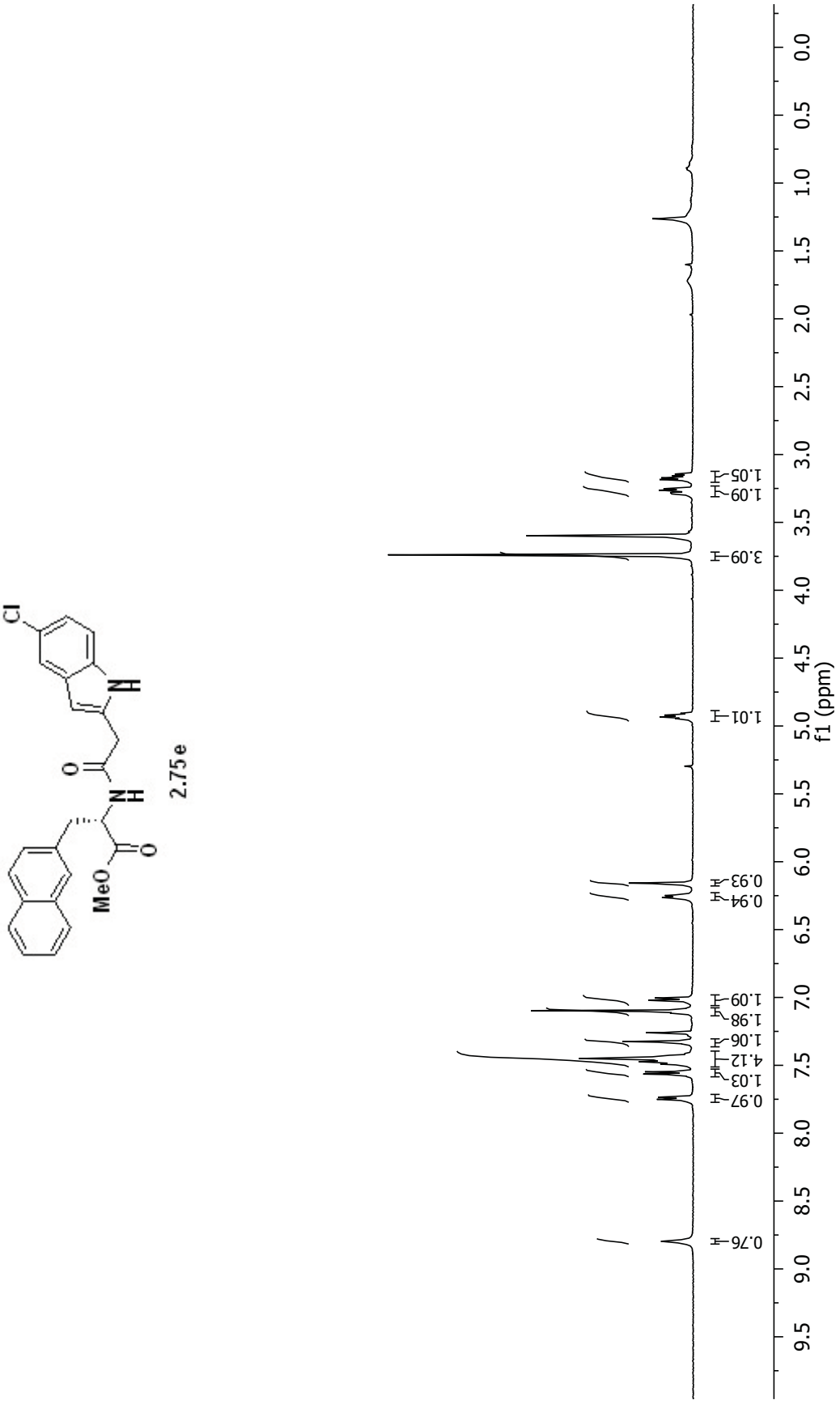

Cl
O
MeO
N
H
O
N
Boc

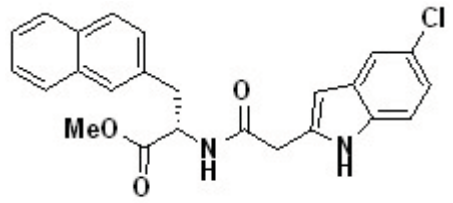
2.76c


200
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
-10
f1 (ppm)

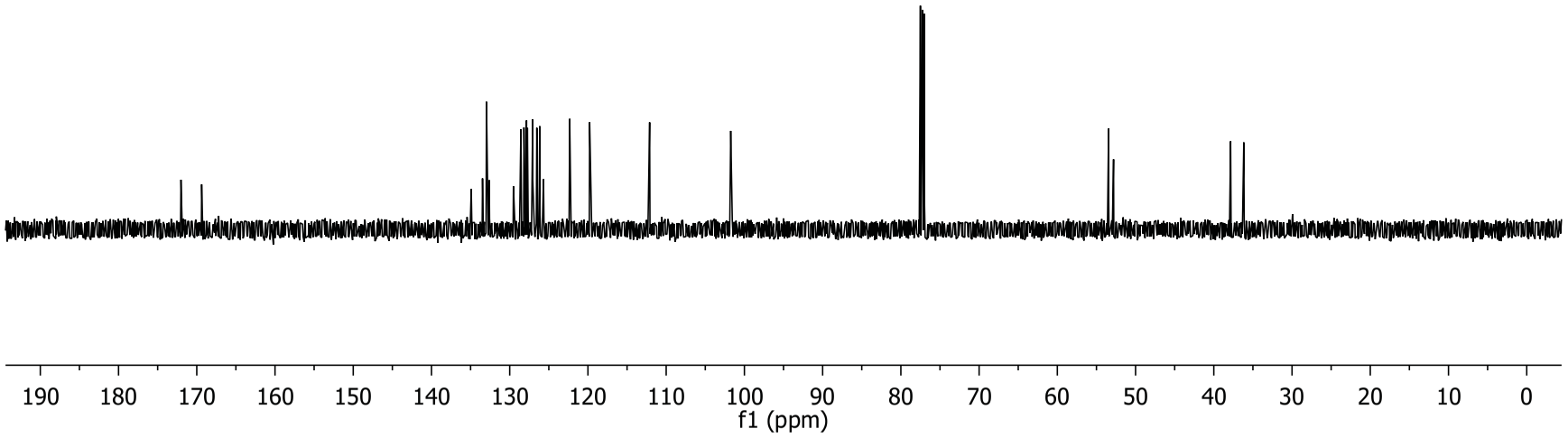

Cl
Boc
N
O
H
N
O
COOtBu
MeO
COOtBu
2.77c
10.09
20.00
5.13
9.64
1.12
1.00
1.92
0.91
0.84
3.00
1.86
0.90
0.90
0.89
0.87
2.87
0.39
0.87
0.87
f1 (ppm)

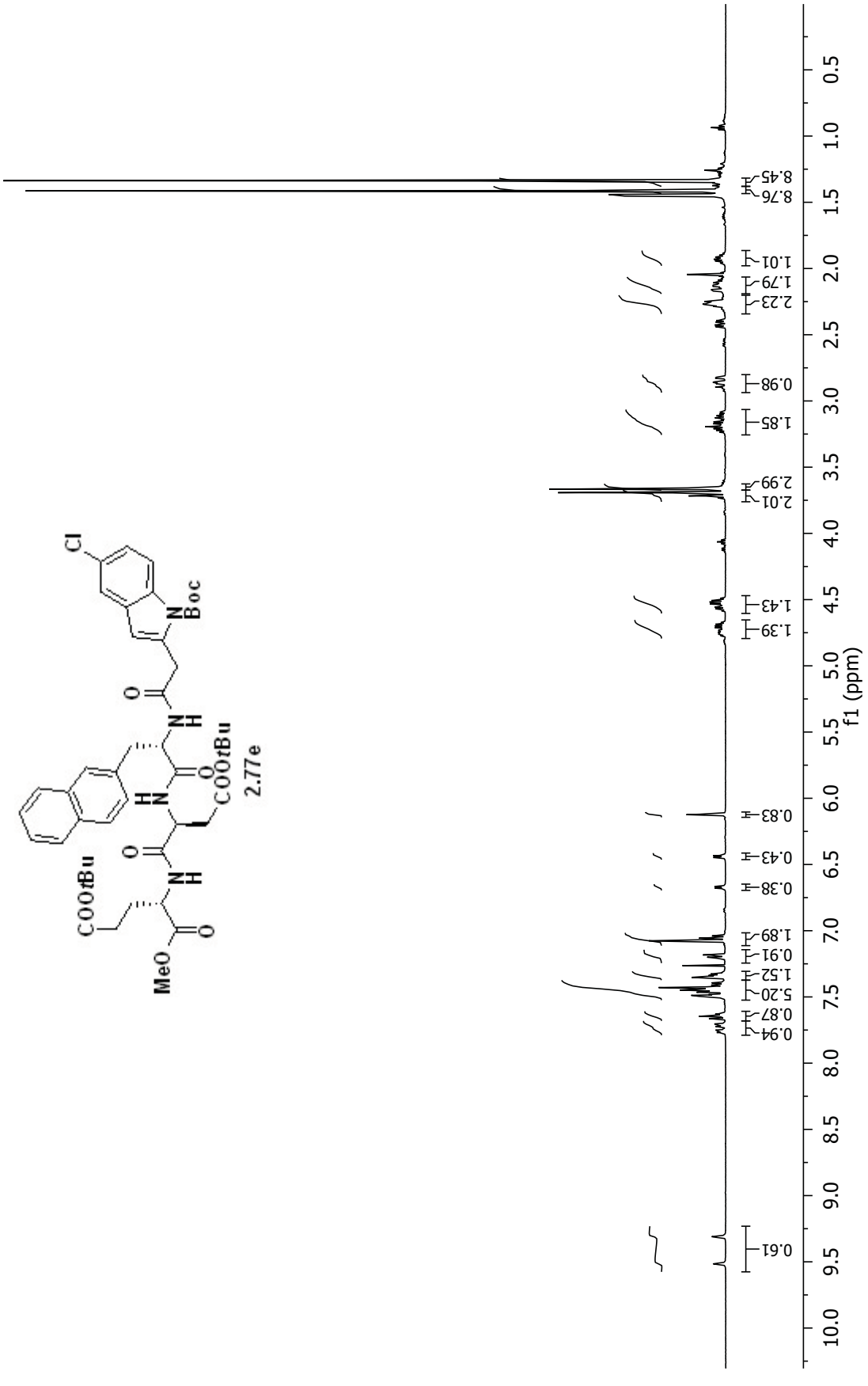

Cl
N
Boc
O
N
H
O
N
H
O
COOtBu
COOtBu
O
N
H
O
MeO
O
2.77c
f1 (ppm)

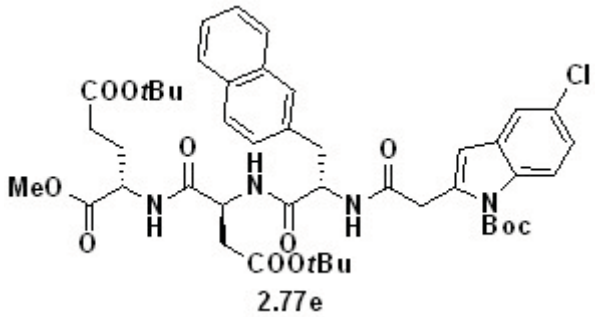

Cl
N
H
O
N
H
MeO
O
2.75e
0.76
0.97
1.03
4.12
1.06
1.98
1.09
0.94
0.93
1.01
3.09
1.09
1.05
9.5
9.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
f1 (ppm)


Cl
O
MeO
N
H
O
N
H

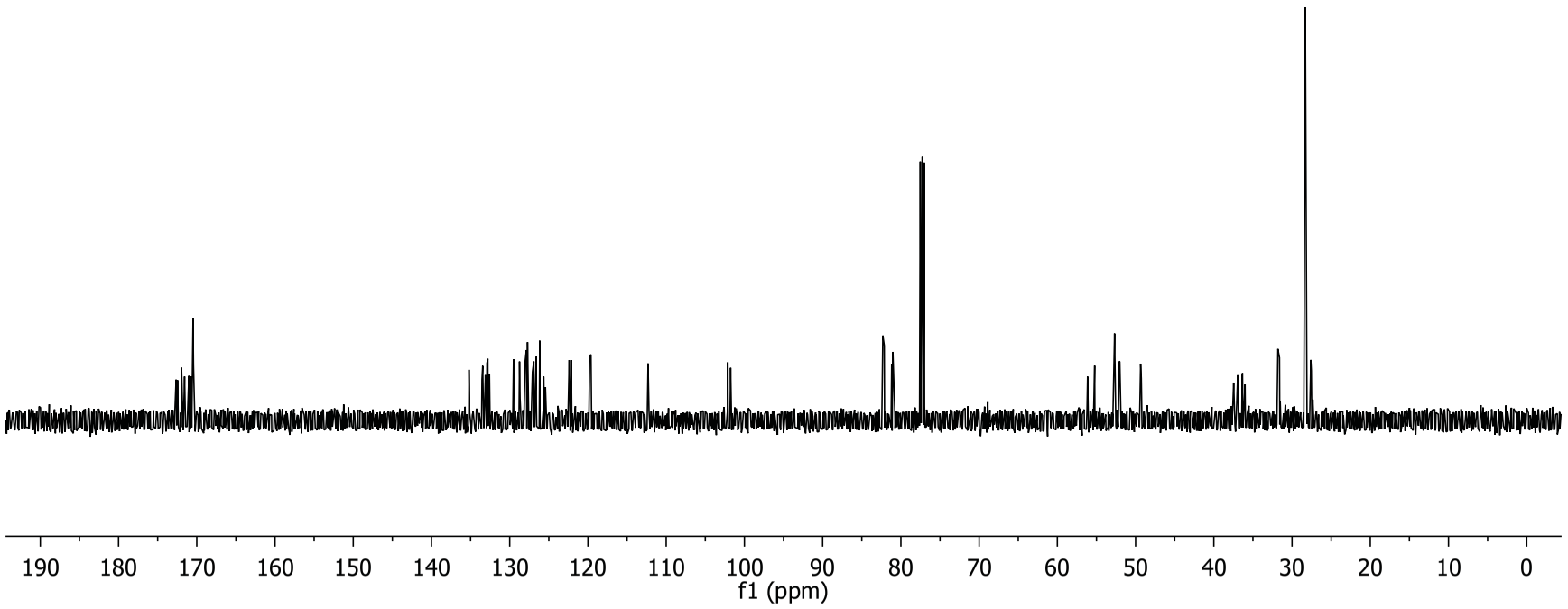
2.75e

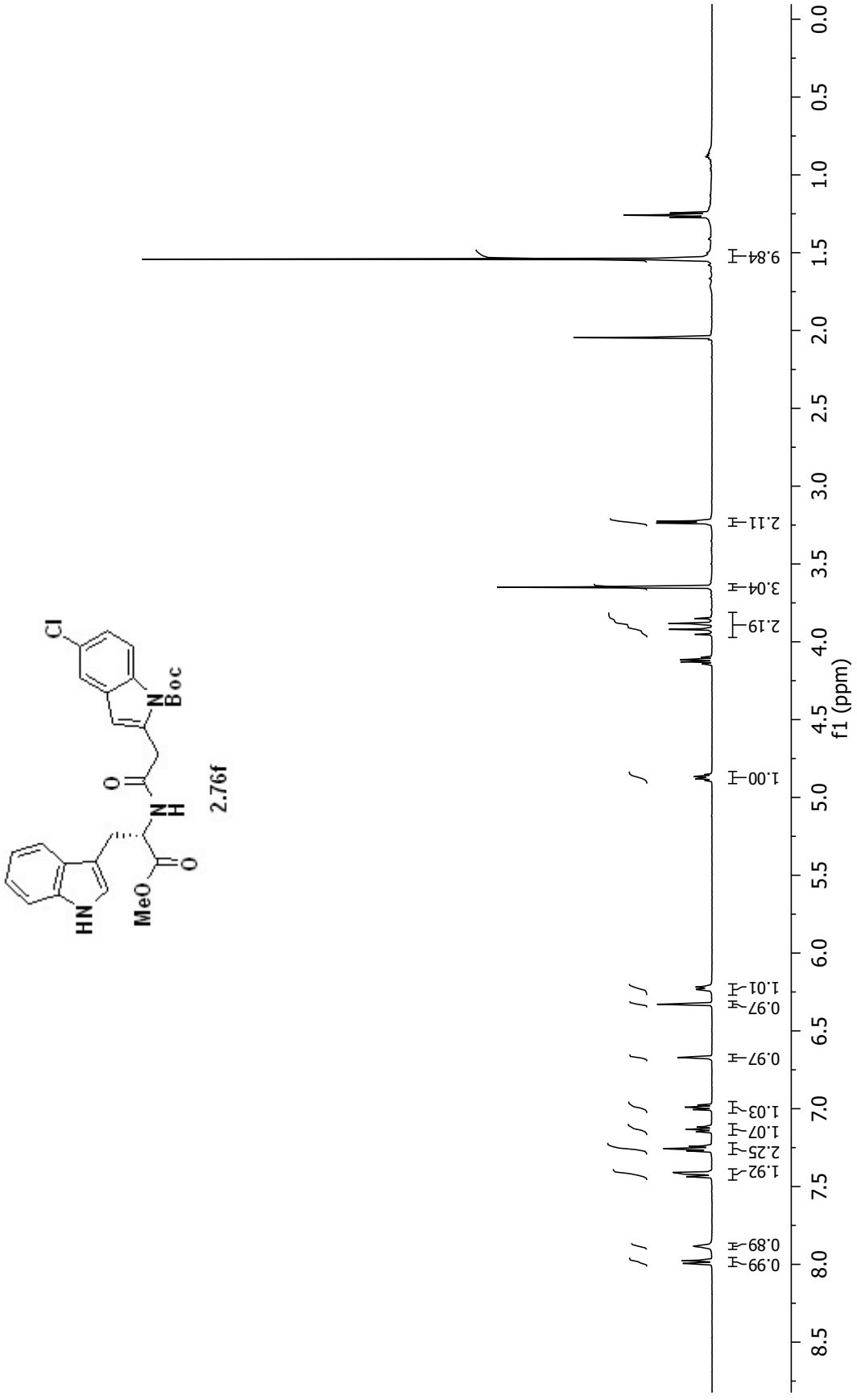

190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
f1 (ppm)

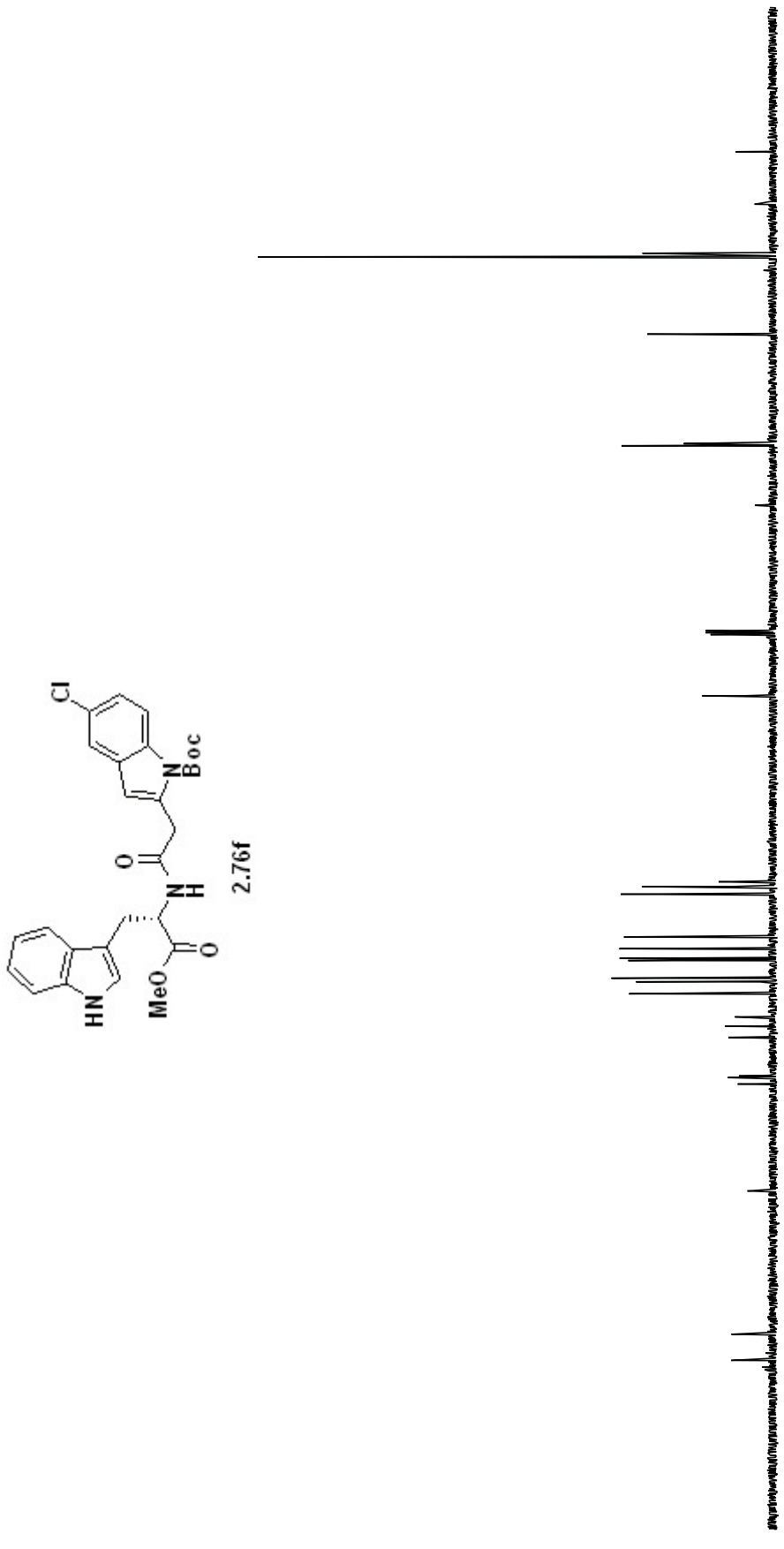

Cl
N
Boc
O
N
H
O
N
H
O
COOtBu
COOtBu
O
N
H
O
MeO
O
2.77e
10.0
9.5
9.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
f1 (ppm)
0.61
0.94
0.87
5.20
1.52
0.91
1.89
0.38
0.43
0.83
1.39
1.43
2.01
2.99
1.85
0.98
2.23
1.79
1.01
8.76
8.45

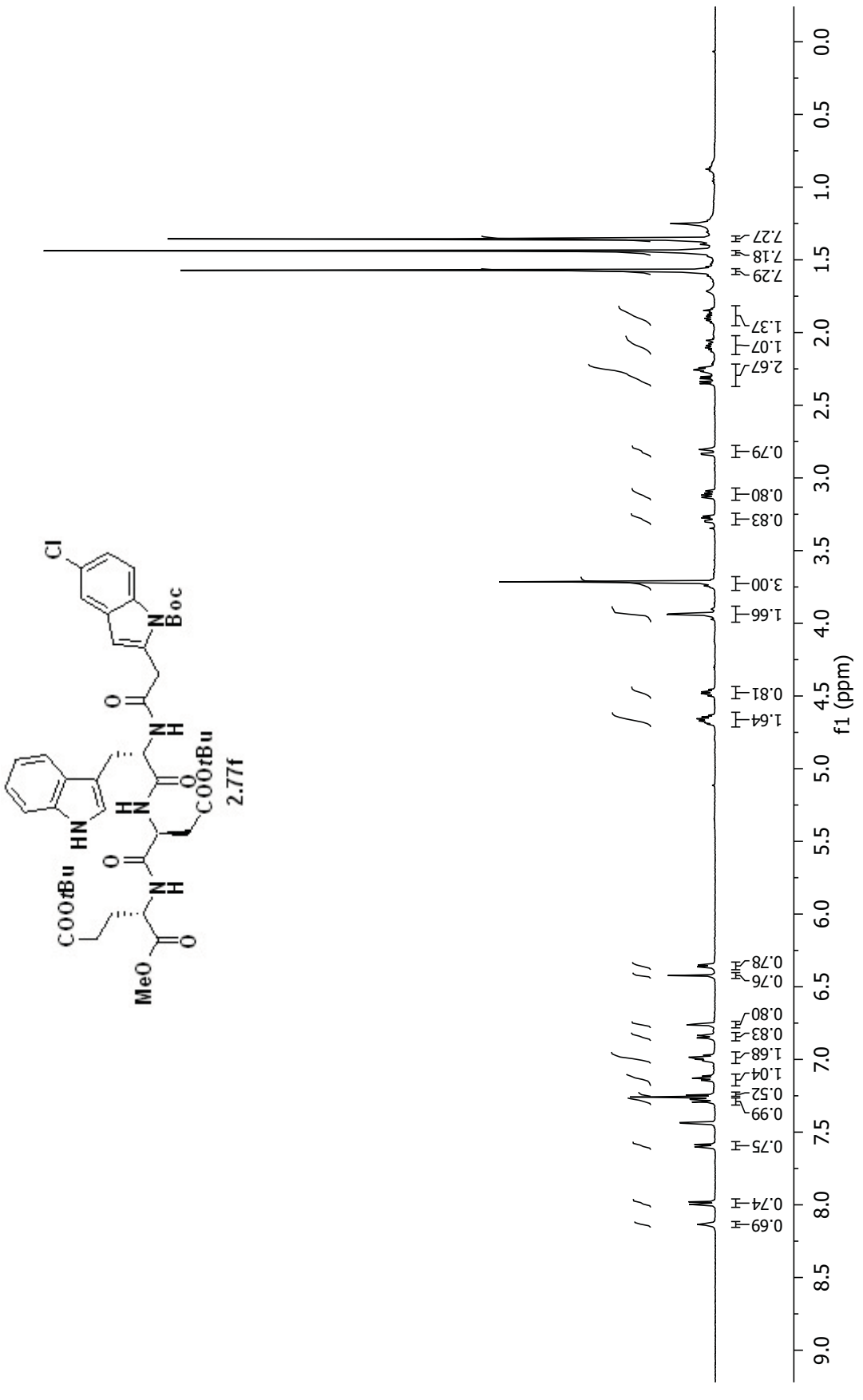

COOtBu
MeO
O
NH
O
HN
O
NH
O
Cl
N
Boc
COOtBu

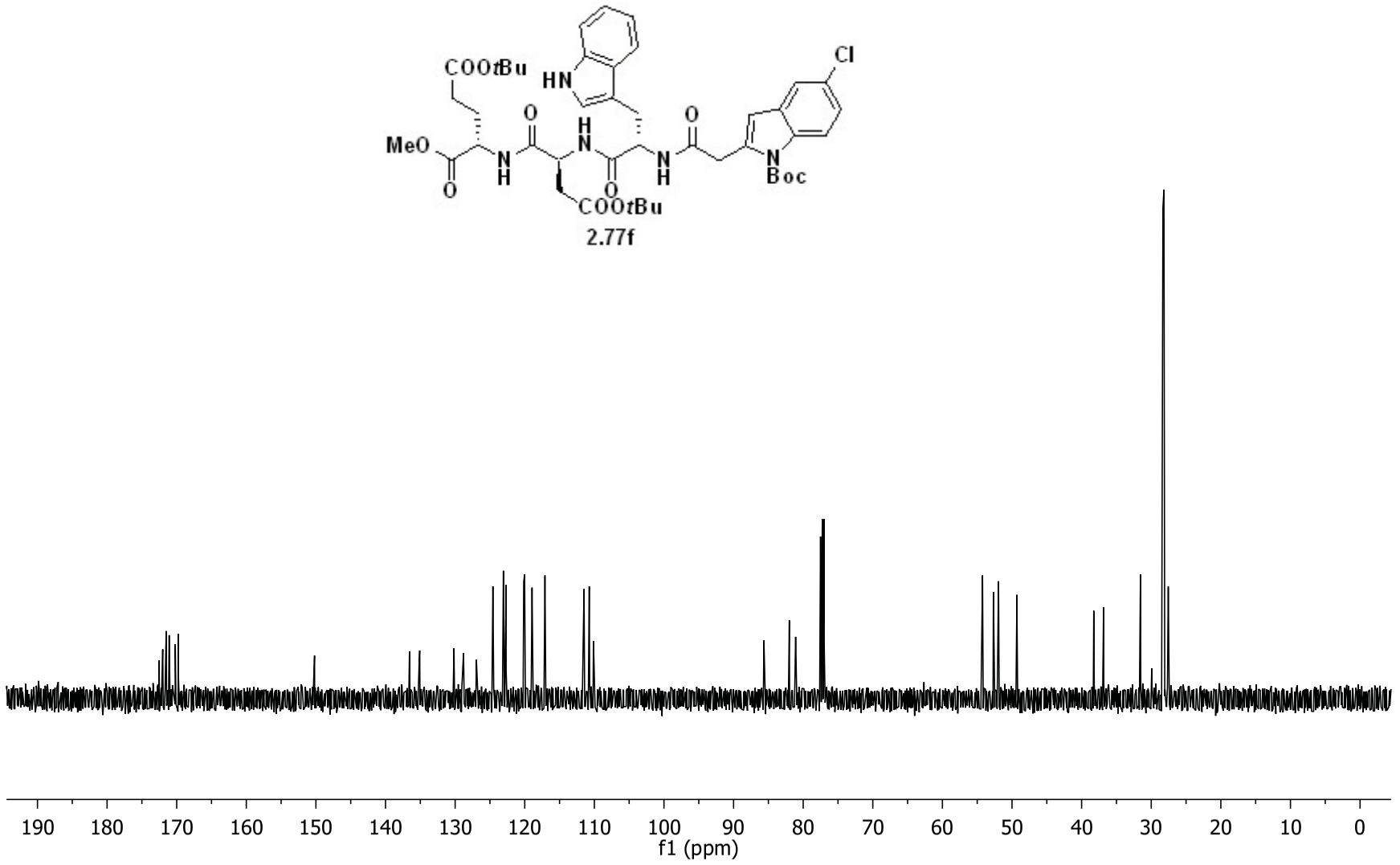
2.77e

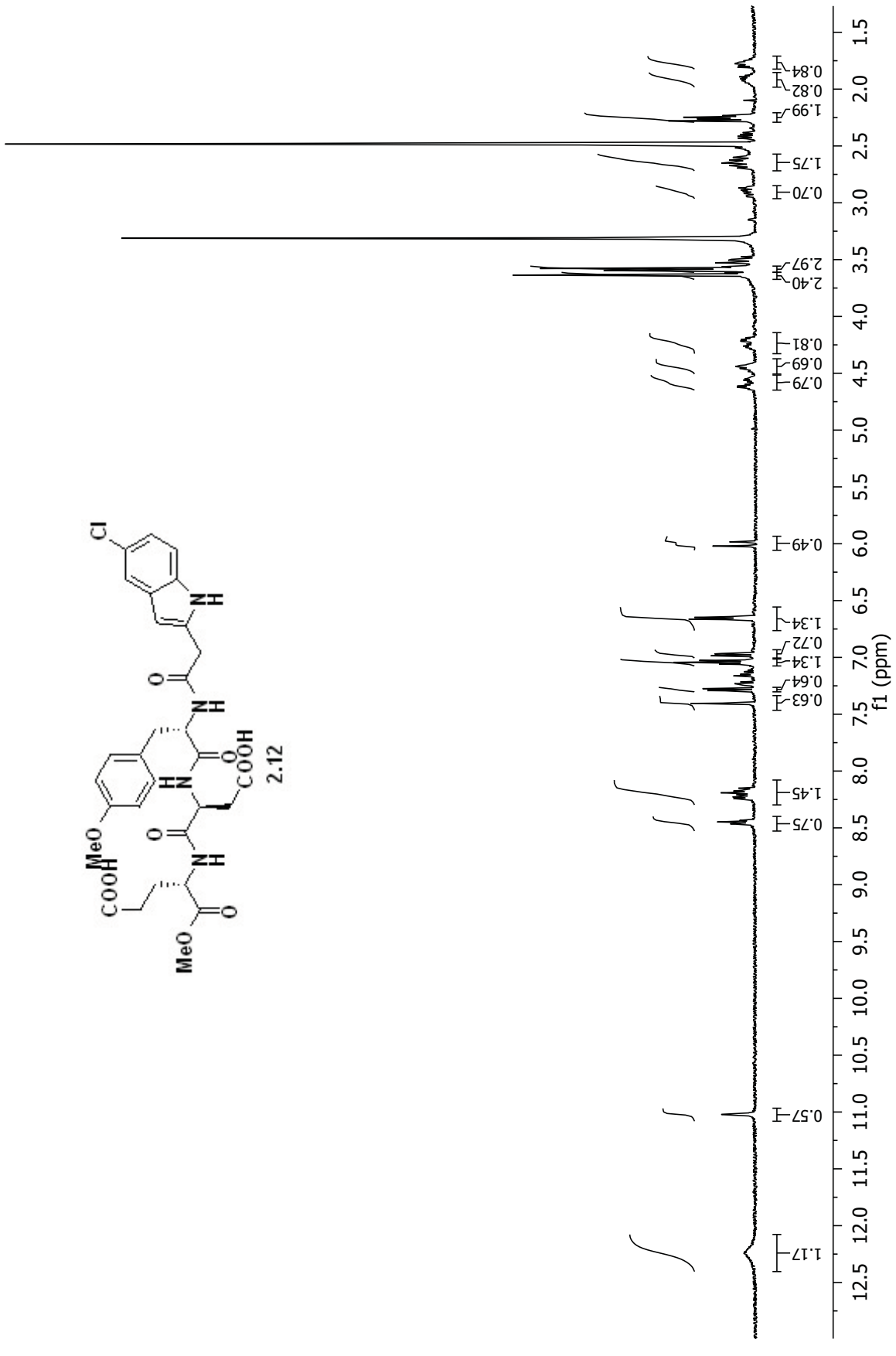

190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
f1 (ppm)

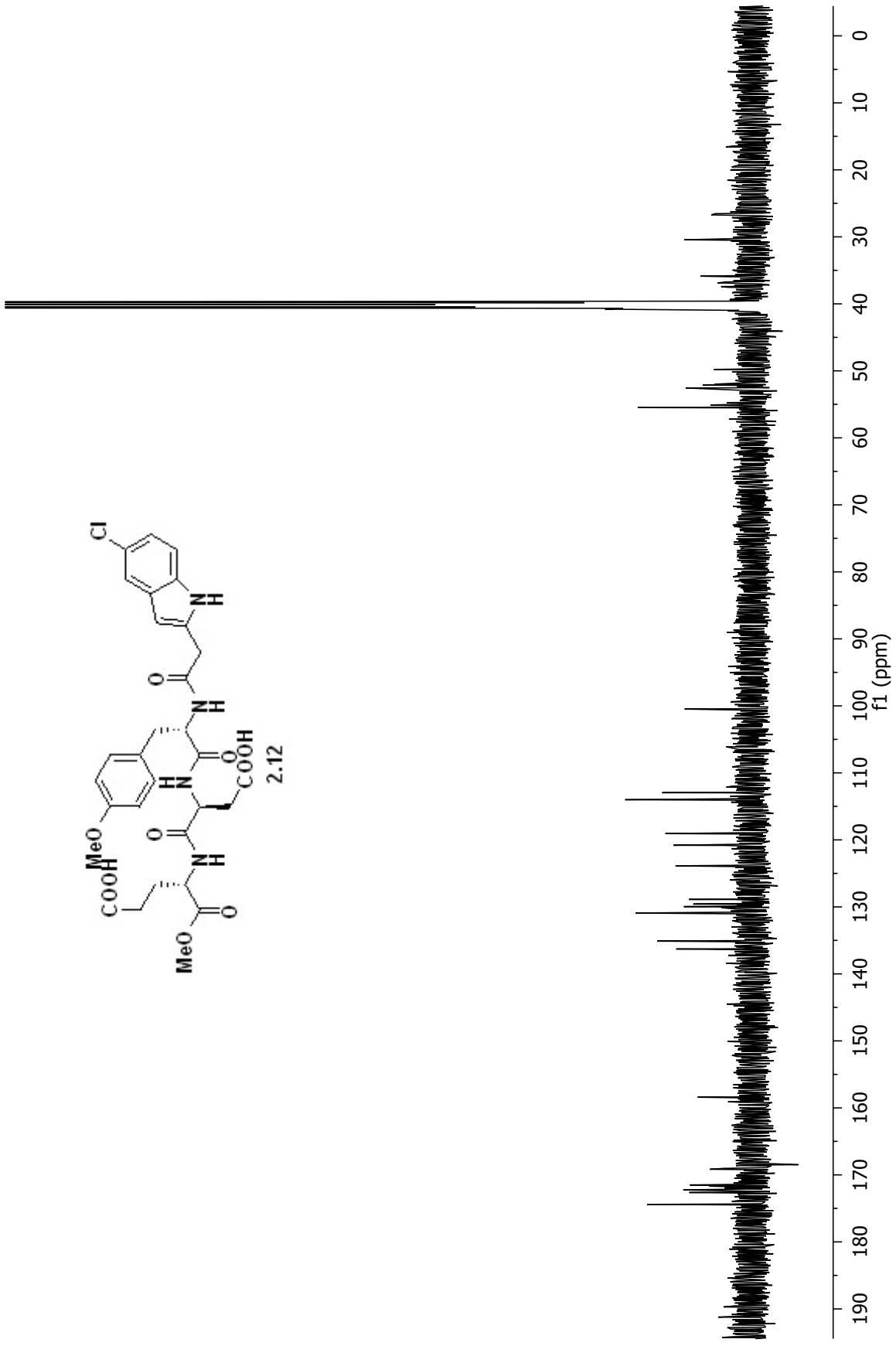

Cl
N
Boc
O
N
H
O
MeO
O
HN
2.76f
9.84
2.11
3.04
2.19
1.00
0.97
1.01
0.97
1.03
1.07
2.25
1.92
0.89
0.99
f1 (ppm)


Cl
N
Boc
O
N
H
O
MeO
O
HN
2.76f
f1 (ppm)

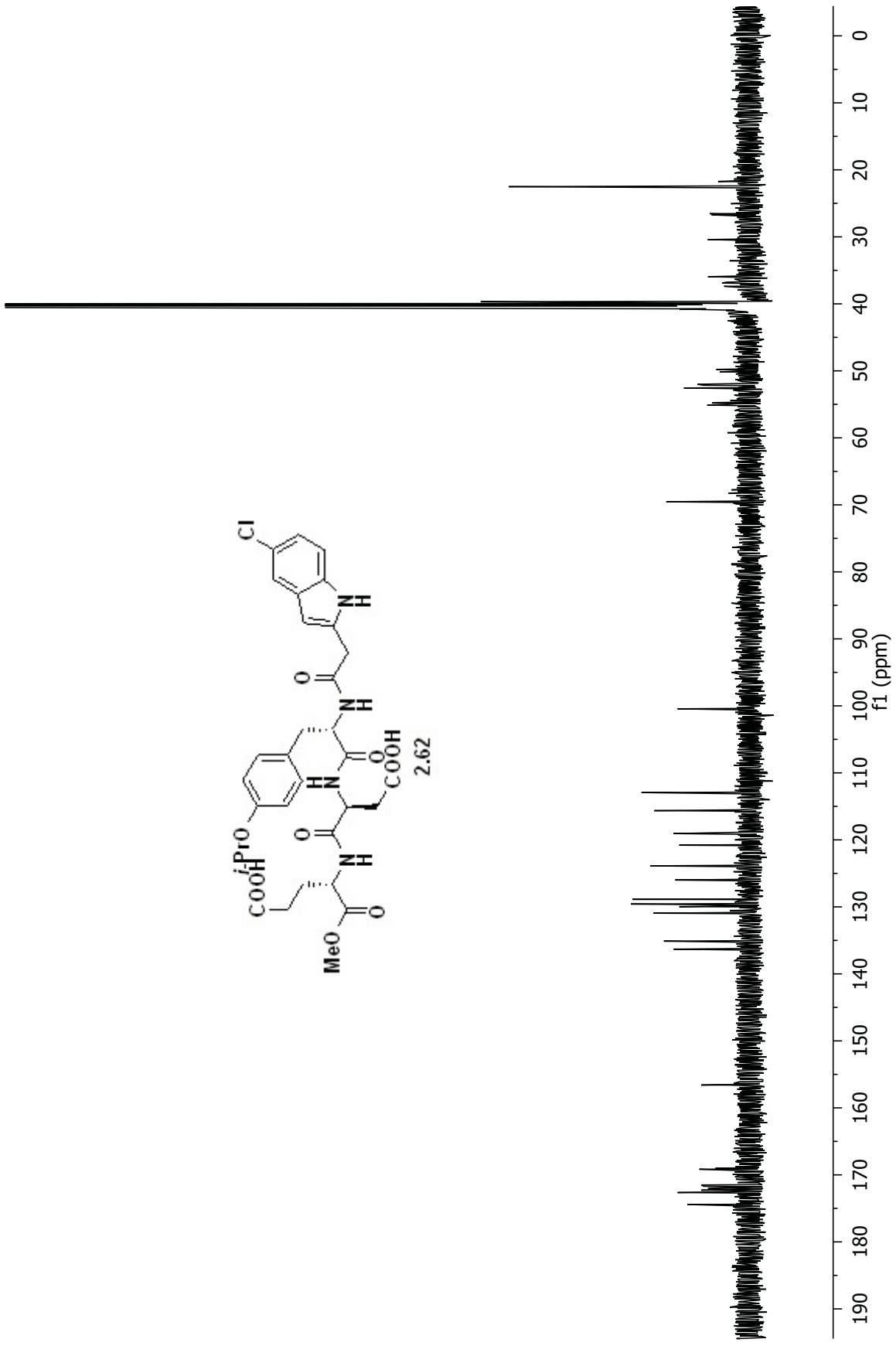

Cl
Boc
N
O
H
N
O
COOtBu
HN
COOtBu
MeO
2.77f
0.69
0.74
0.75
0.99
0.52
1.04
1.68
0.83
0.80
0.76
0.78
1.64
0.81
1.66
3.00
0.83
0.80
0.79
2.67
1.07
1.37
7.29
7.18
7.27
9.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
f1 (ppm)

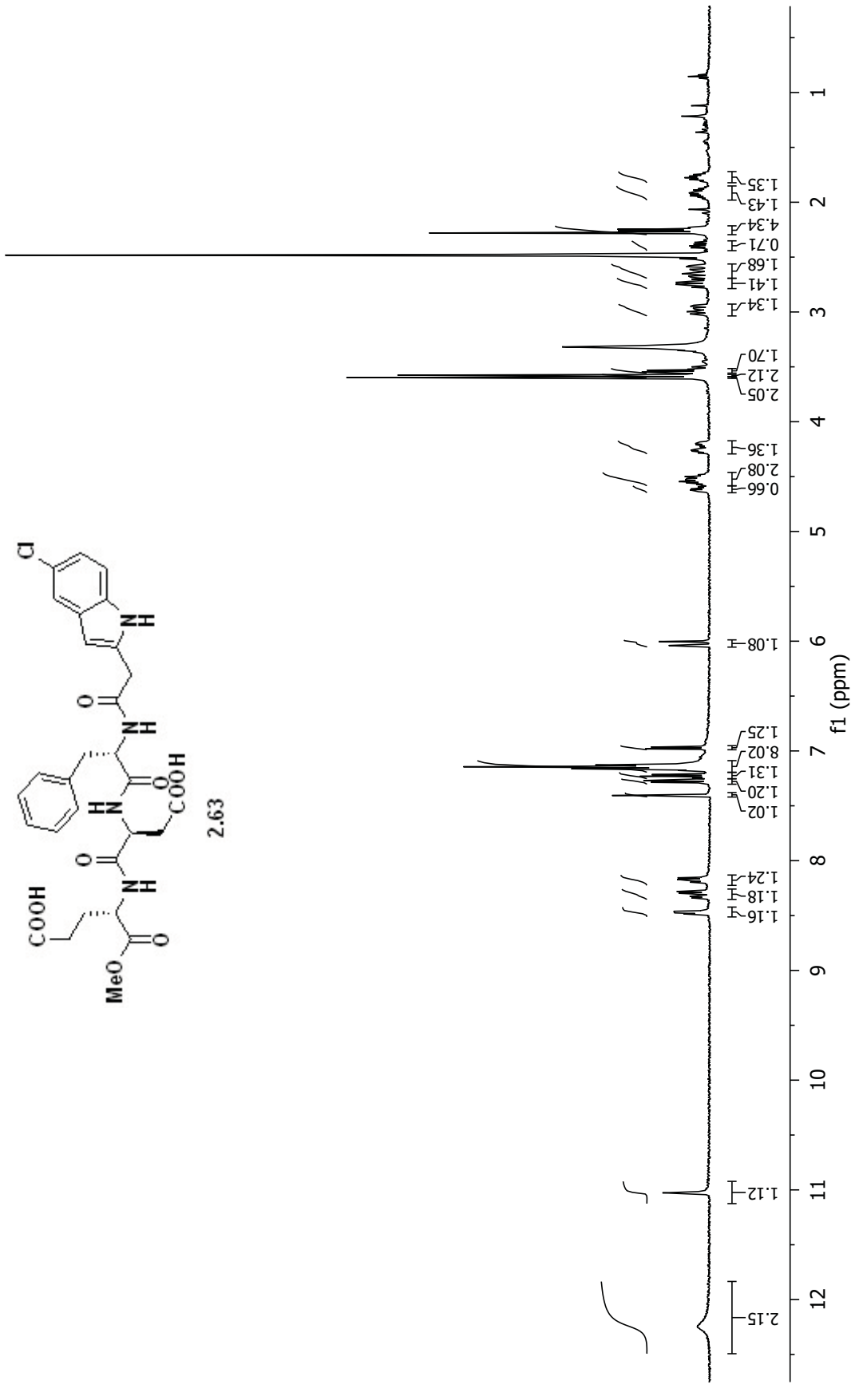

COOtBu
HN
Cl
MeO
O
N
H
O
H
N
O
N
H
N
Boc
COOtBu
2.77f
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
f1 (ppm)

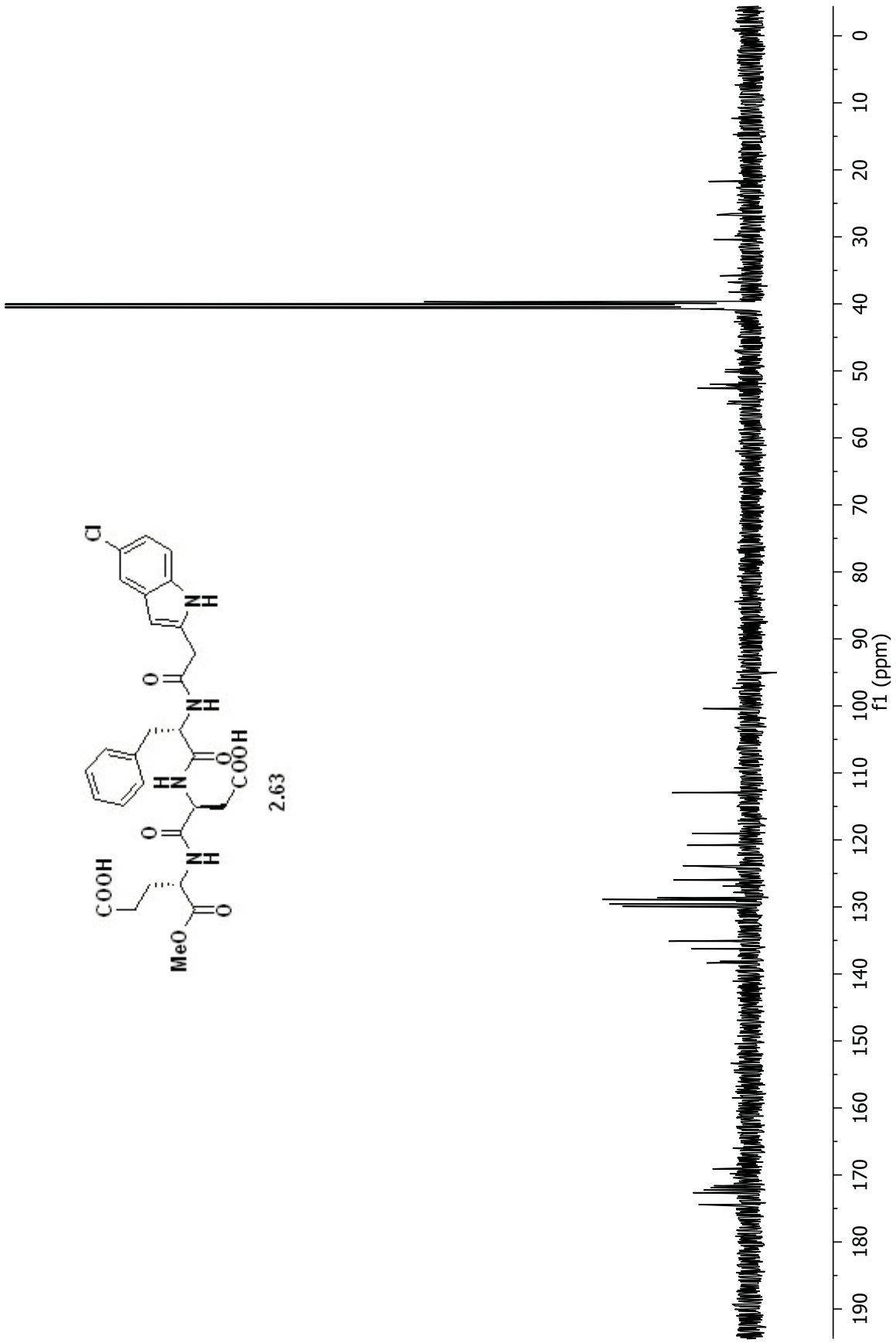

Cl
MeO
COOH
MeO
COOH
2.12
1.17
0.57
0.75
1.45
0.63
0.64
1.34
0.72
1.34
0.49
1.34
0.79
0.69
0.81
2.40
2.97
0.70
1.75
1.99
0.82
0.84
f1 (ppm)

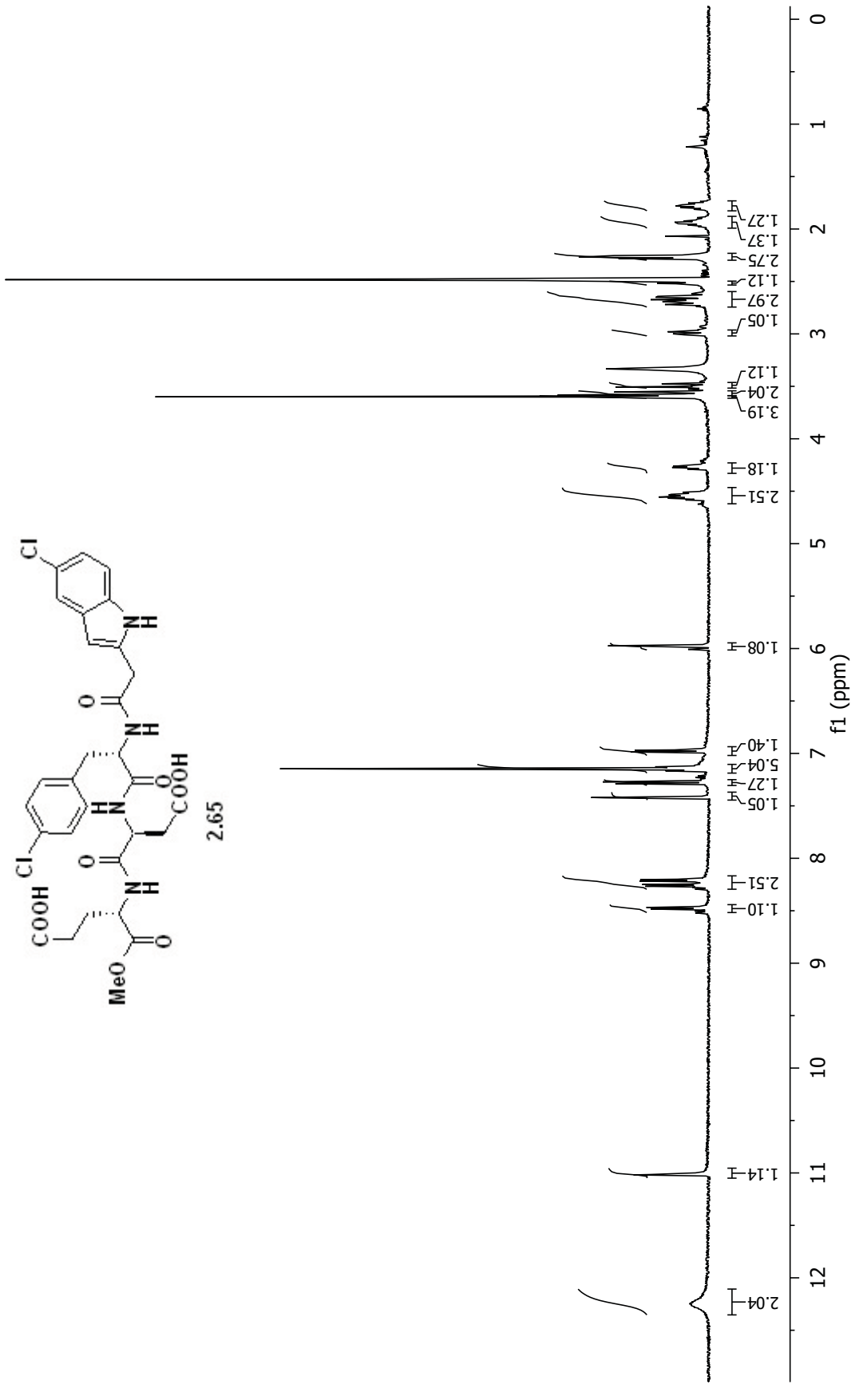

Cl
NH
O
NH
O
COOH
NH
O
MeO
COOH
NH
O
MeO
2.12
0
10
20
30
40
50
60
70
80
90
100
110
120
130
140
150
160
170
180
190
f1 (ppm)

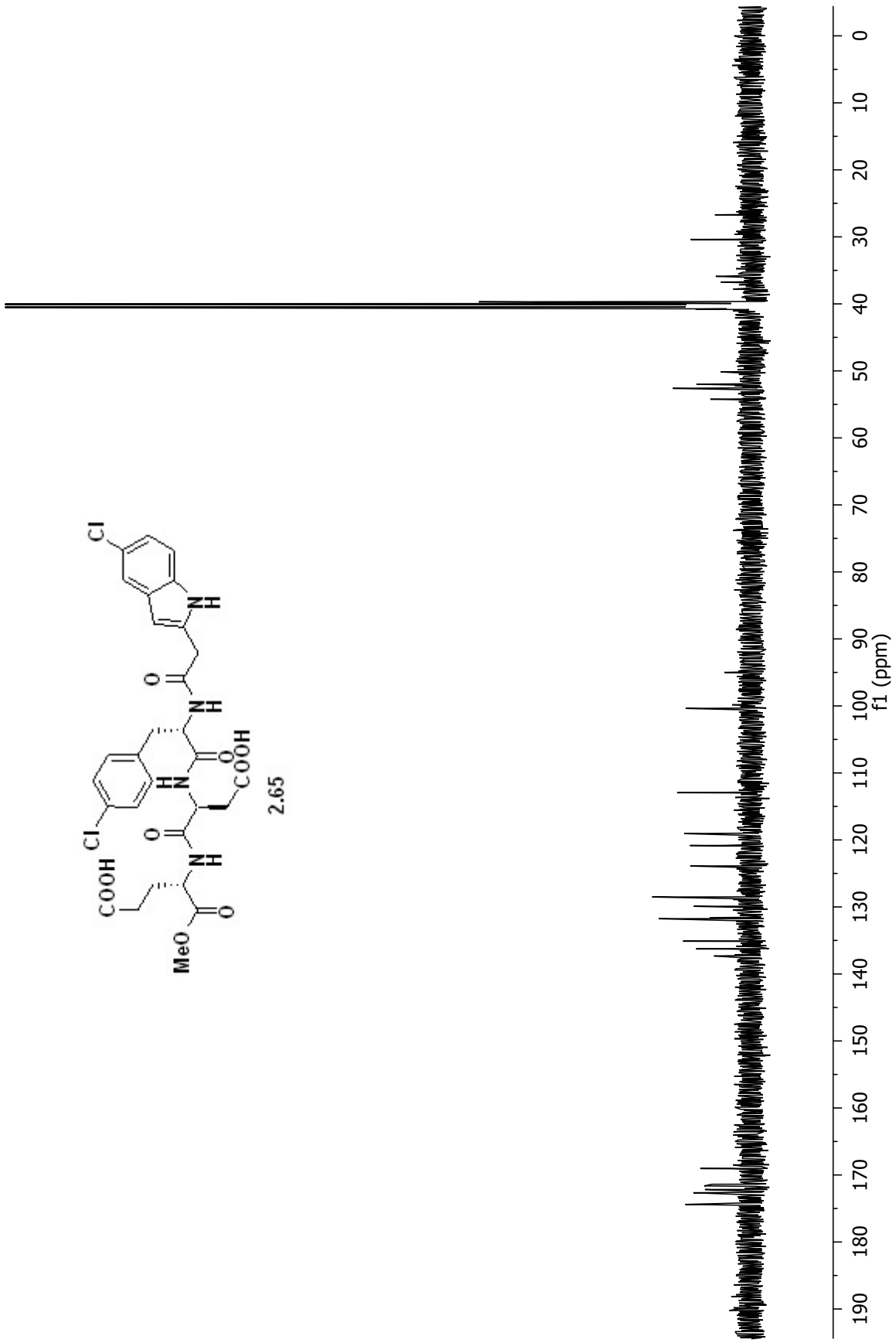
Cl
N
H
O
N
H
O
O
COOH
2.62
N
H
O
O
MeO
O
COOiPr
10.42
2.65
2.39
5.79
1.53
0.40
3.79
1.75
2.09
3.77
3.00
2.29
3.48
0.94
0.95
1.42
3.38
1.85
3.05
1.40
0.94
1.56
1.31
1.70
1.72
1.60
1.11
2.14
f1 (ppm)

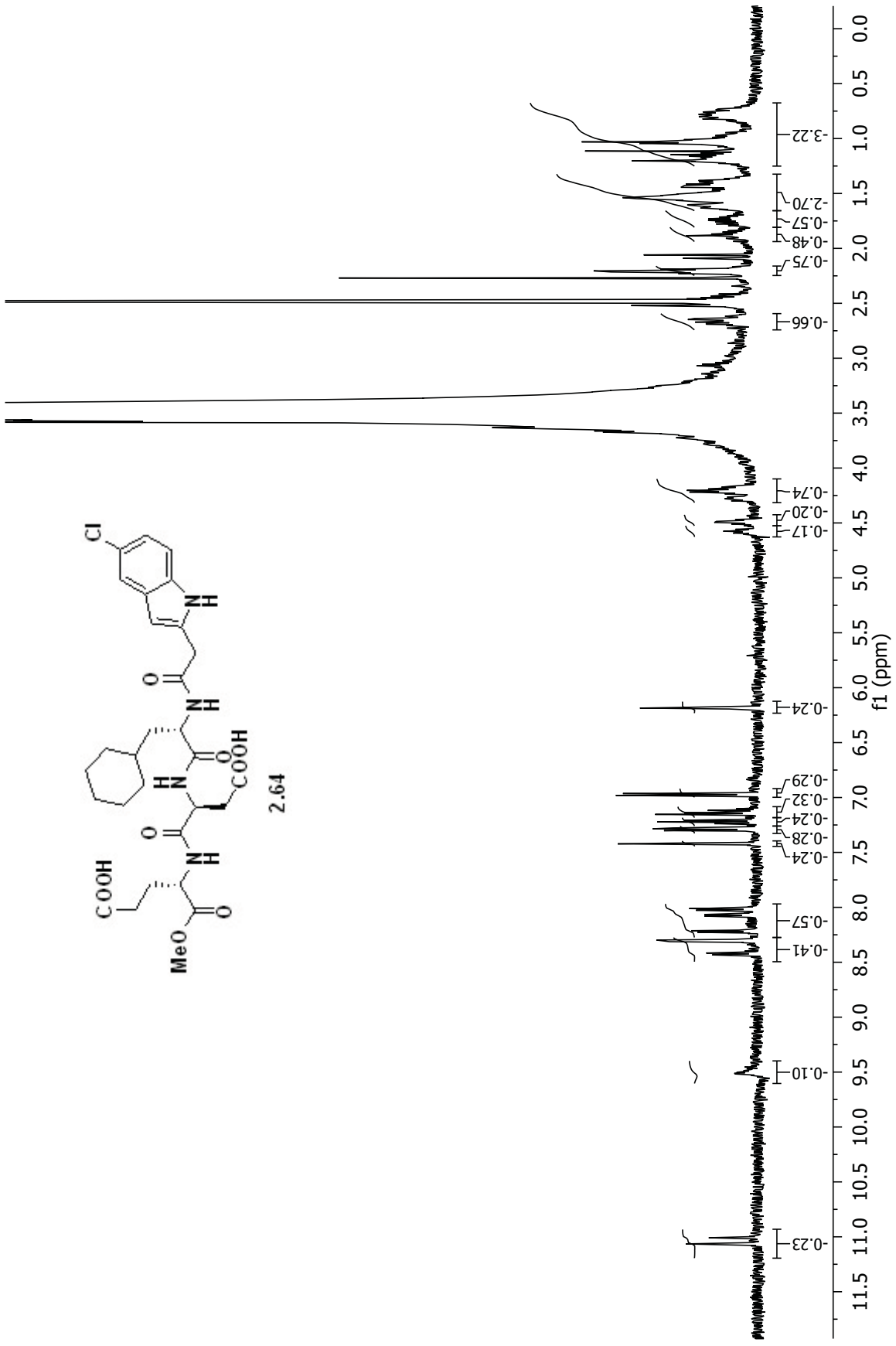

Cl
NH
O
NH
O
COOH
2.62
NH
O
COOiPr
MeO
O
f1 (ppm)
0
10
20
30
40
50
60
70
80
90
100
110
120
130
140
150
160
170
180
190

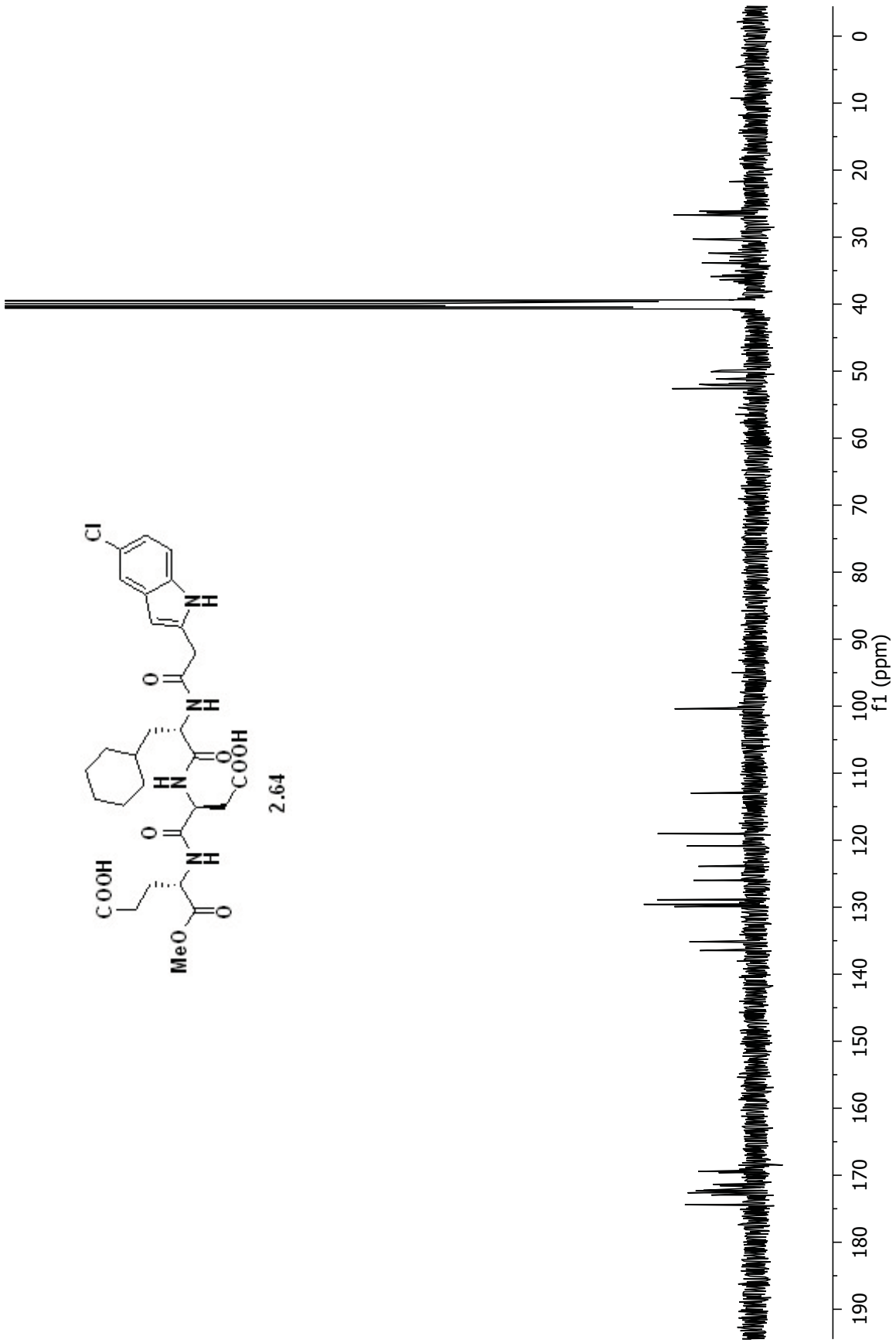

Cl
COOH
MeO
COOH
2.63
2.15
1.12
1.16
1.18
1.24
1.02
1.20
1.31
8.02
1.25
1.08
0.66
2.08
1.36
2.05
2.12
1.70
1.34
1.41
1.68
0.71
4.34
1.43
1.35
f1 (ppm)

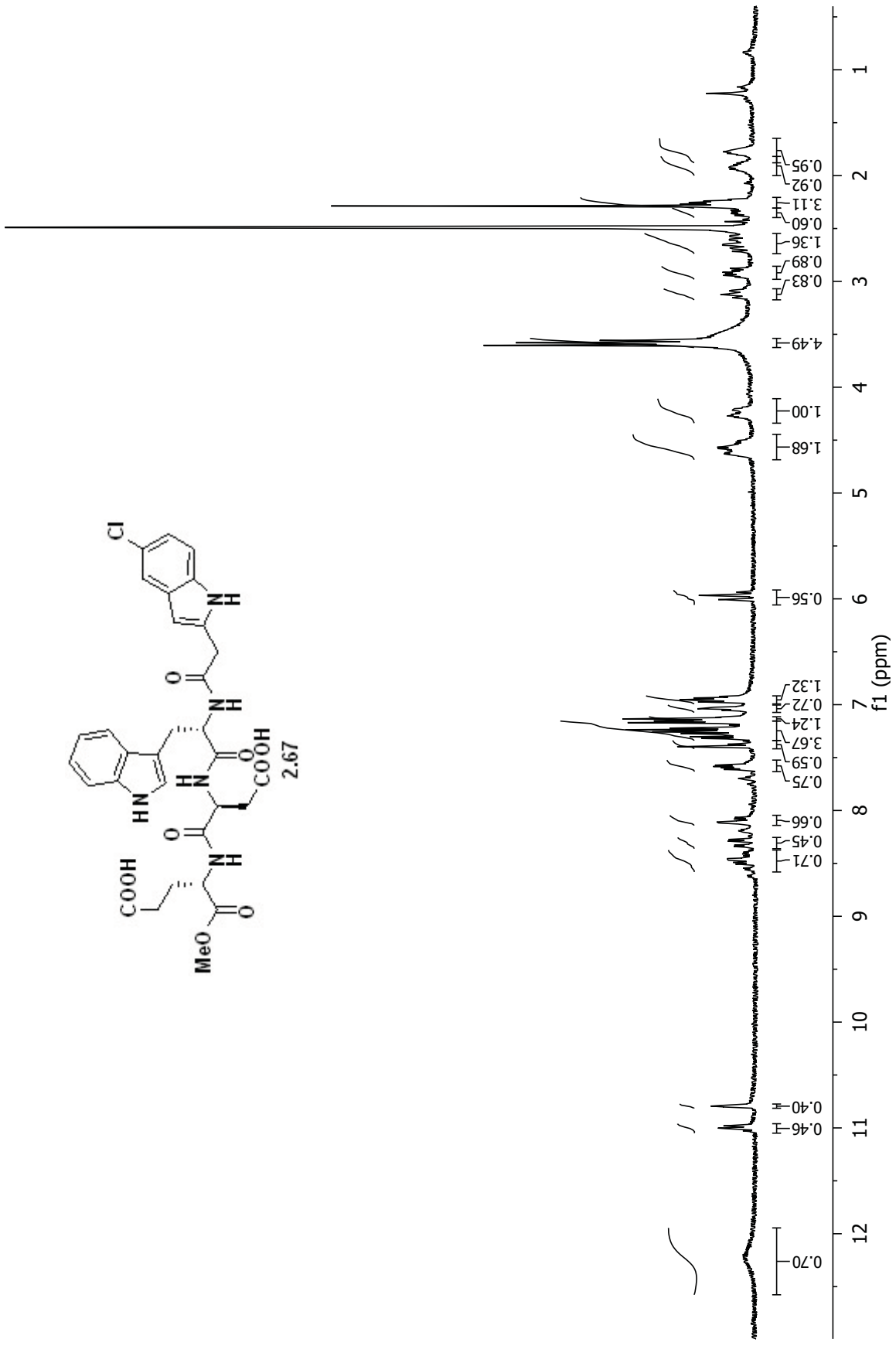

Cl
NH
O
NH
O
NH
O
COOH
O
NH
O
COOH
MeO
O
2.63
0
10
20
30
40
50
60
70
80
90
100
110
120
130
140
150
160
170
180
190
f1 (ppm)

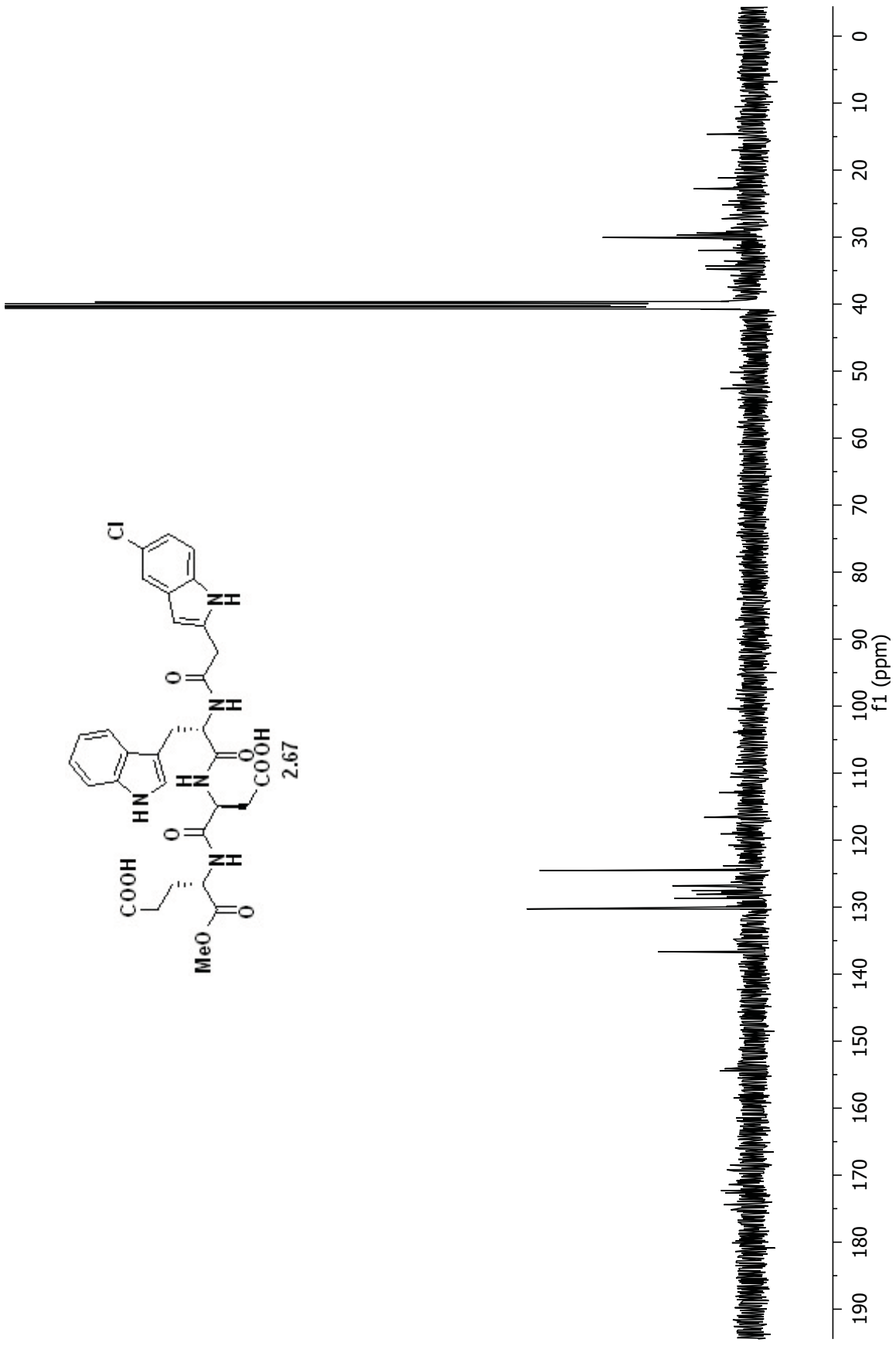

Cl
COOH
MeO
Cl
COOH
2.65
2.04
1.14
1.10
2.51
1.05
1.27
5.04
1.40
1.08
2.51
1.18
3.19
2.04
1.12
1.05
2.97
1.12
2.75
1.37
1.27
f1 (ppm)

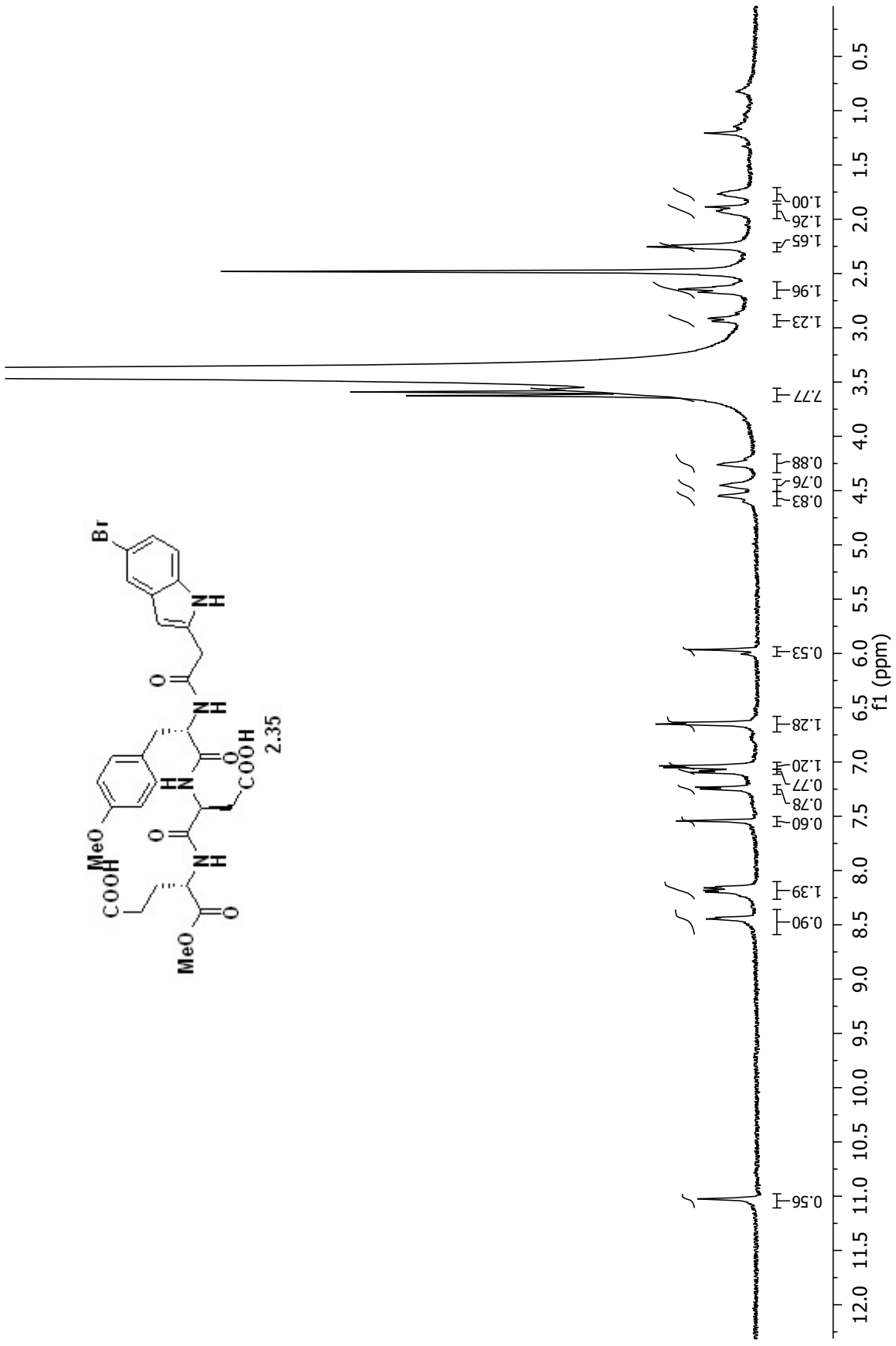

Cl
NH
O
NH
O
COOH
H
N
O
Cl
COOH
NH
O
MeO
2.65
0
10
20
30
40
50
60
70
80
90
100
110
120
130
140
150
160
170
180
190
f1 (ppm)

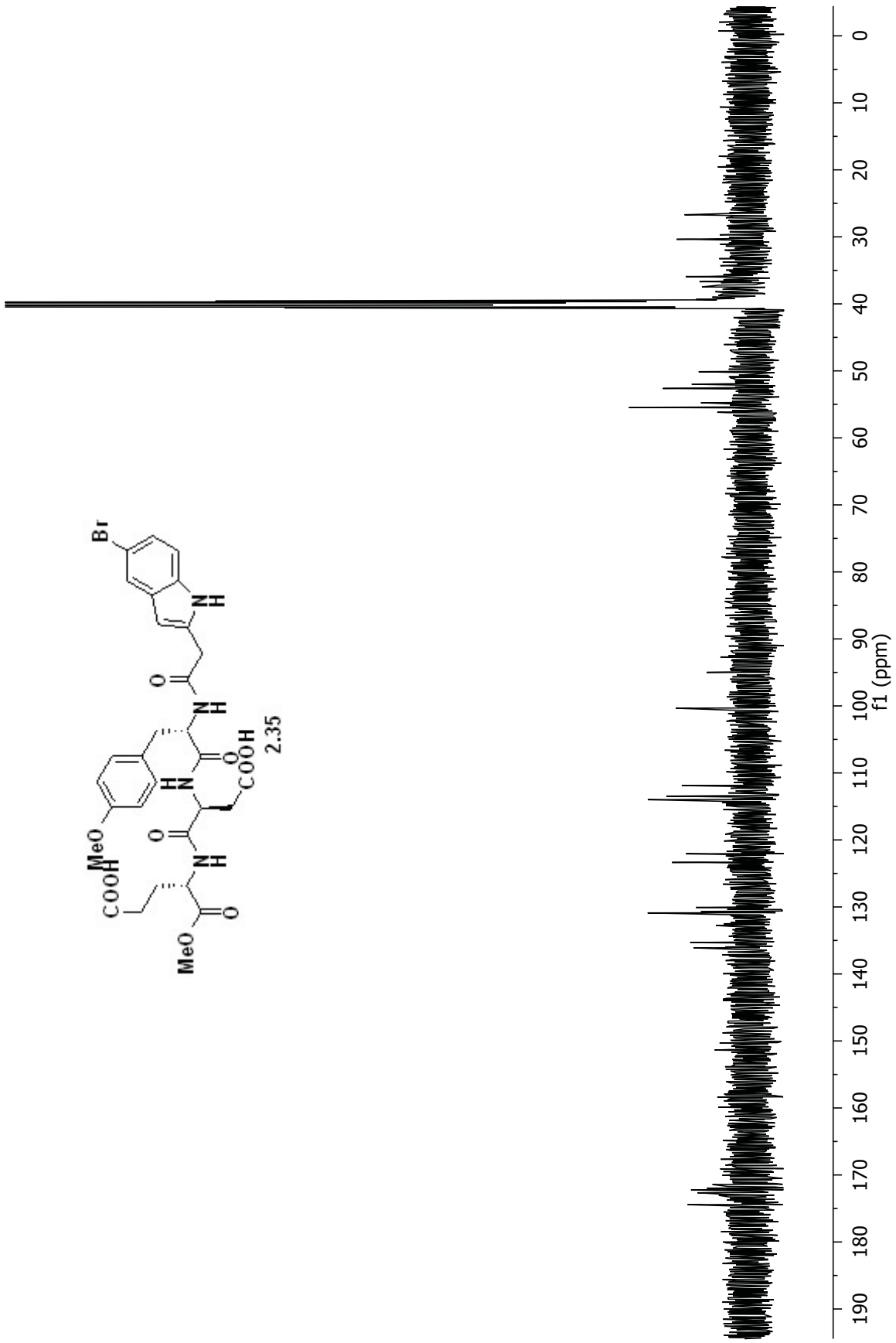

Cl
COOH
MeO
COOH
2.64
0.23
0.10
0.41
0.57
0.24
0.28
0.24
0.32
0.29
0.24
0.74
0.20
0.17
0.66
0.75
0.48
0.57
2.70
3.22
f1 (ppm)

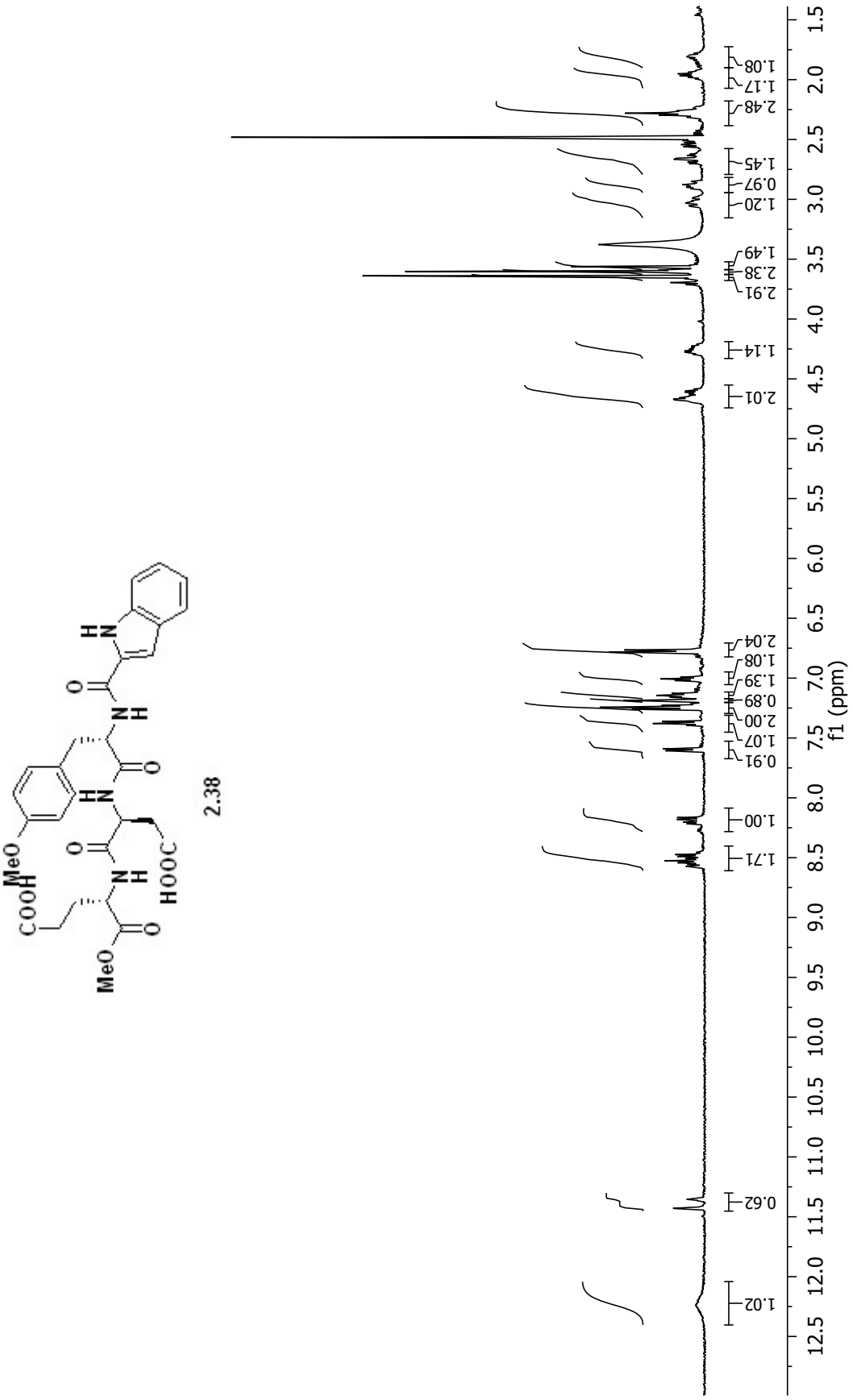

Cl
N
H
O
N
H
O
N
H
O
COOH
COOH
MeO
O
2.64
0
10
20
30
40
50
60
70
80
90
100
110
120
130
140
150
160
170
180
190
f1 (ppm)

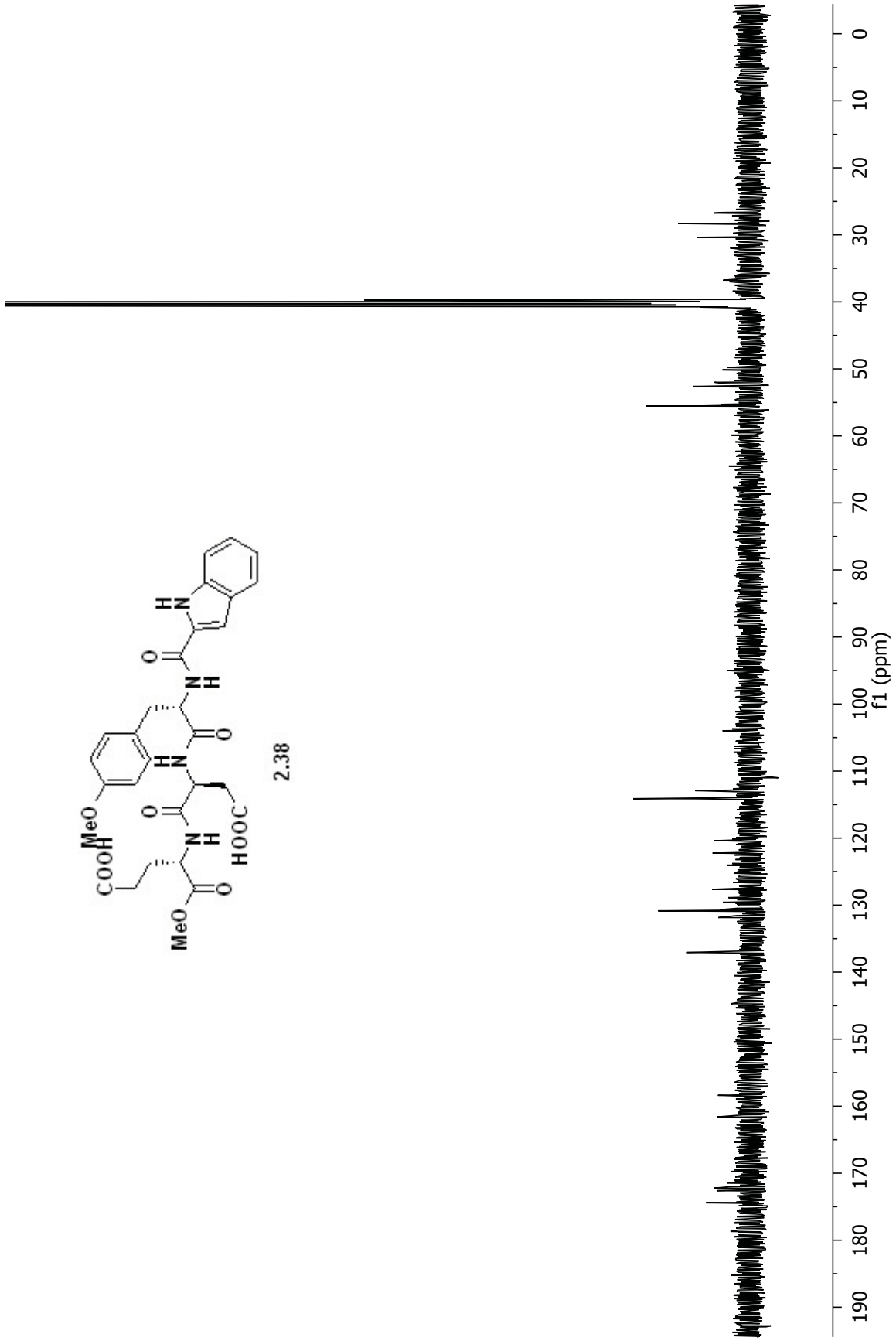

Cl
COOH
MeO
COOH
2.67
0.70
0.46
0.40
0.71
0.45
0.66
0.75
0.59
3.67
1.24
0.72
1.32
0.56
1.68
1.00
4.49
0.83
0.89
1.36
0.60
3.11
0.92
0.95
f1 (ppm)

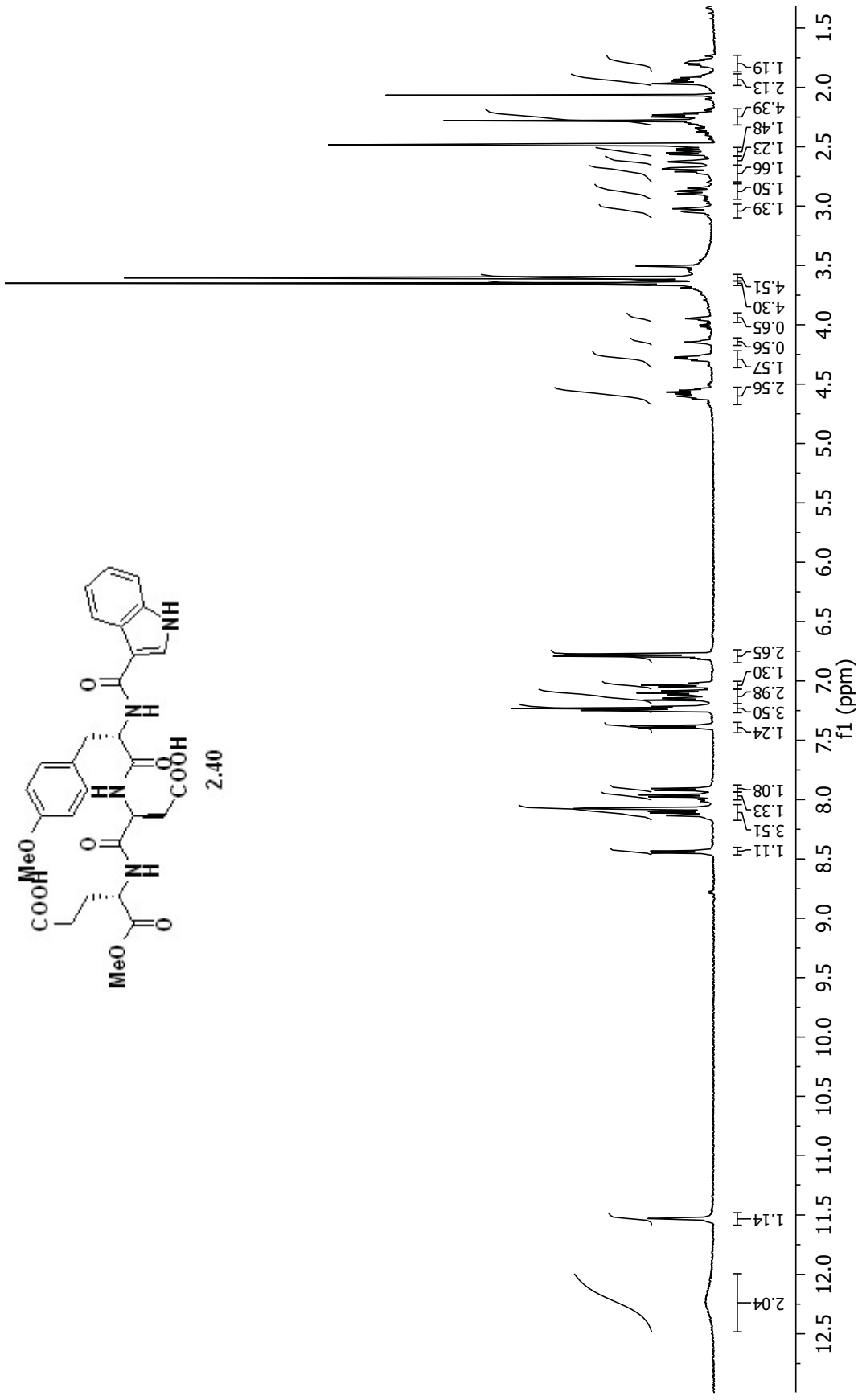

Cl
HN
NH
O
NH
O
NH
O
COOH
2.67
COOH
MeO
O
f1 (ppm)
0
10
20
30
40
50
60
70
80
90
100
110
120
130
140
150
160
170
180
190

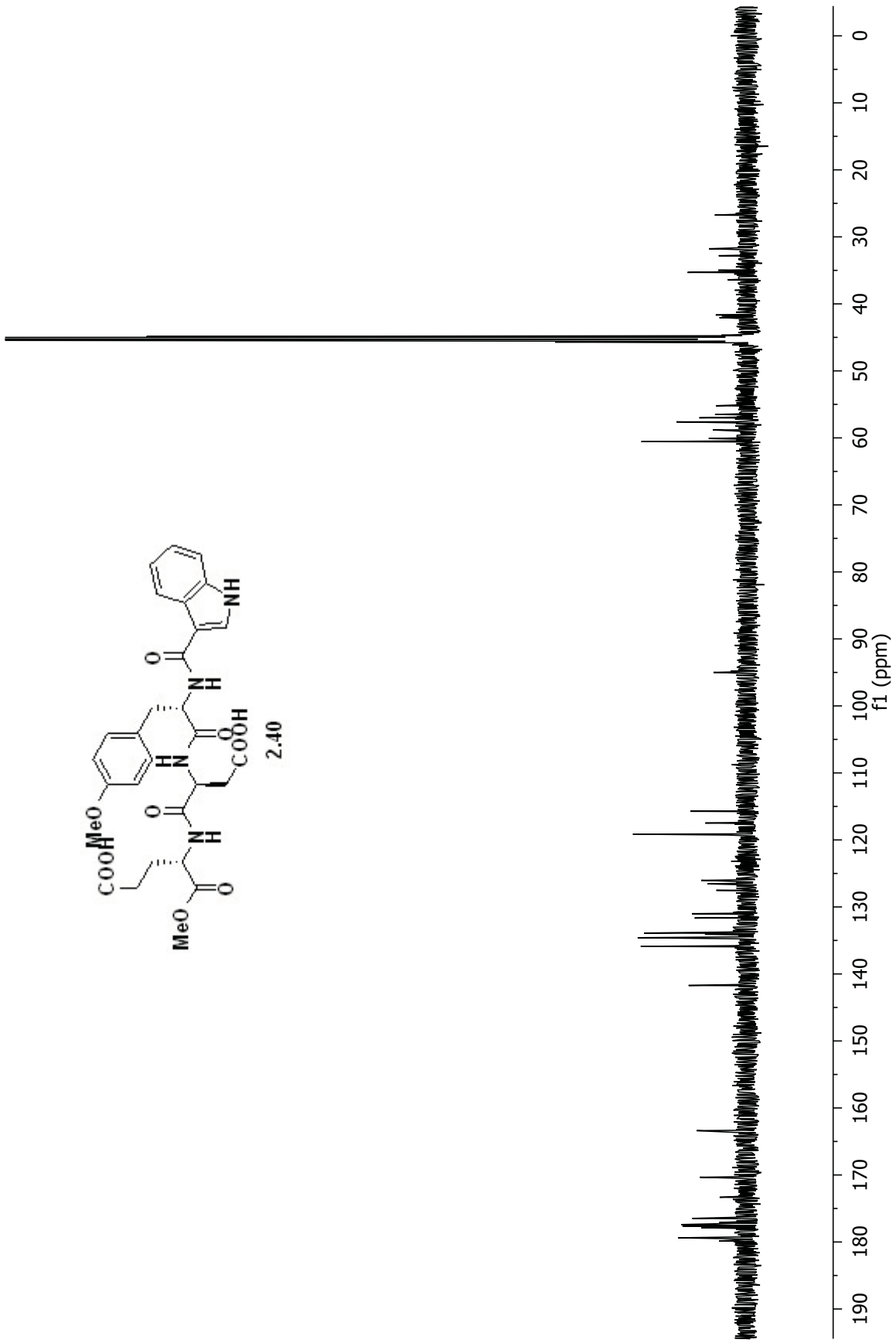

Br
MeO
COOH
MeO
COOH
2.35
0.56
1.39
0.90
0.60
0.78
0.77
1.20
1.28
0.53
0.83
0.76
0.88
7.77
1.23
1.96
1.65
1.26
1.00
f1 (ppm)

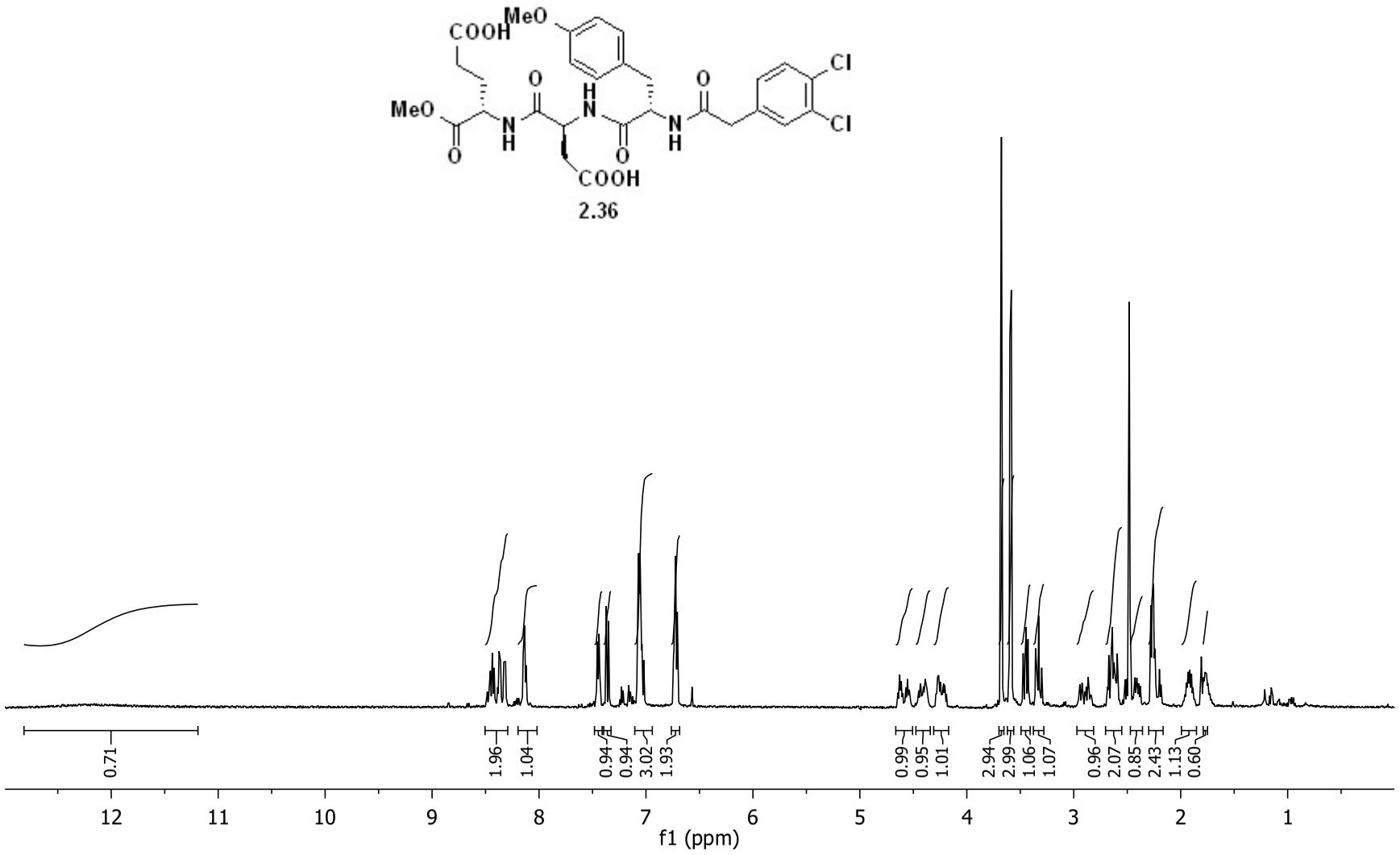

Br
NH
O
NH
O
COOH
2.35
NH
O
MeO
COOH
NH
O
MeO
0
10
20
30
40
50
60
70
80
90
100
110
120
130
140
150
160
170
180
190
f1 (ppm)

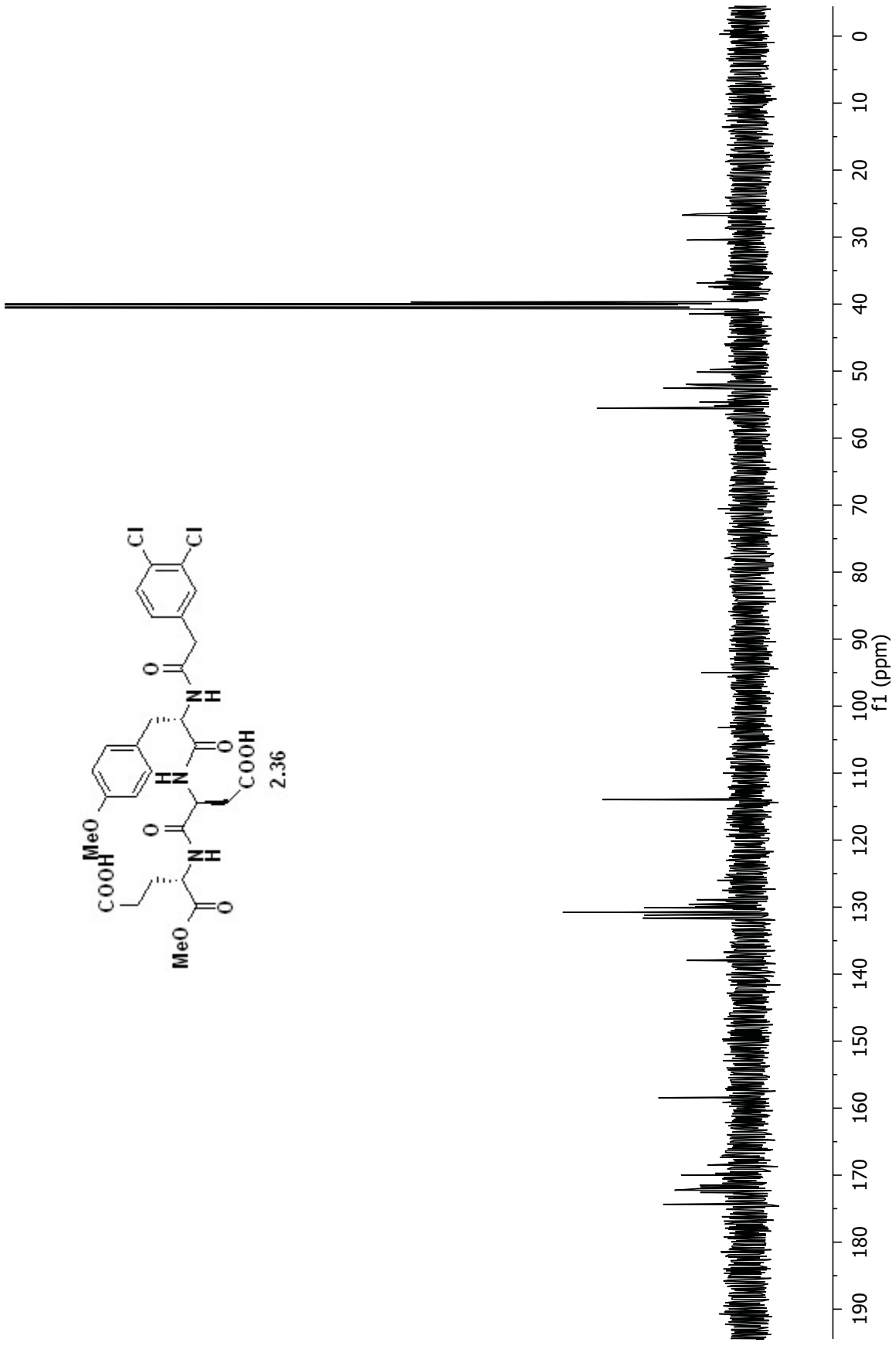

2.38
f1 (ppm)

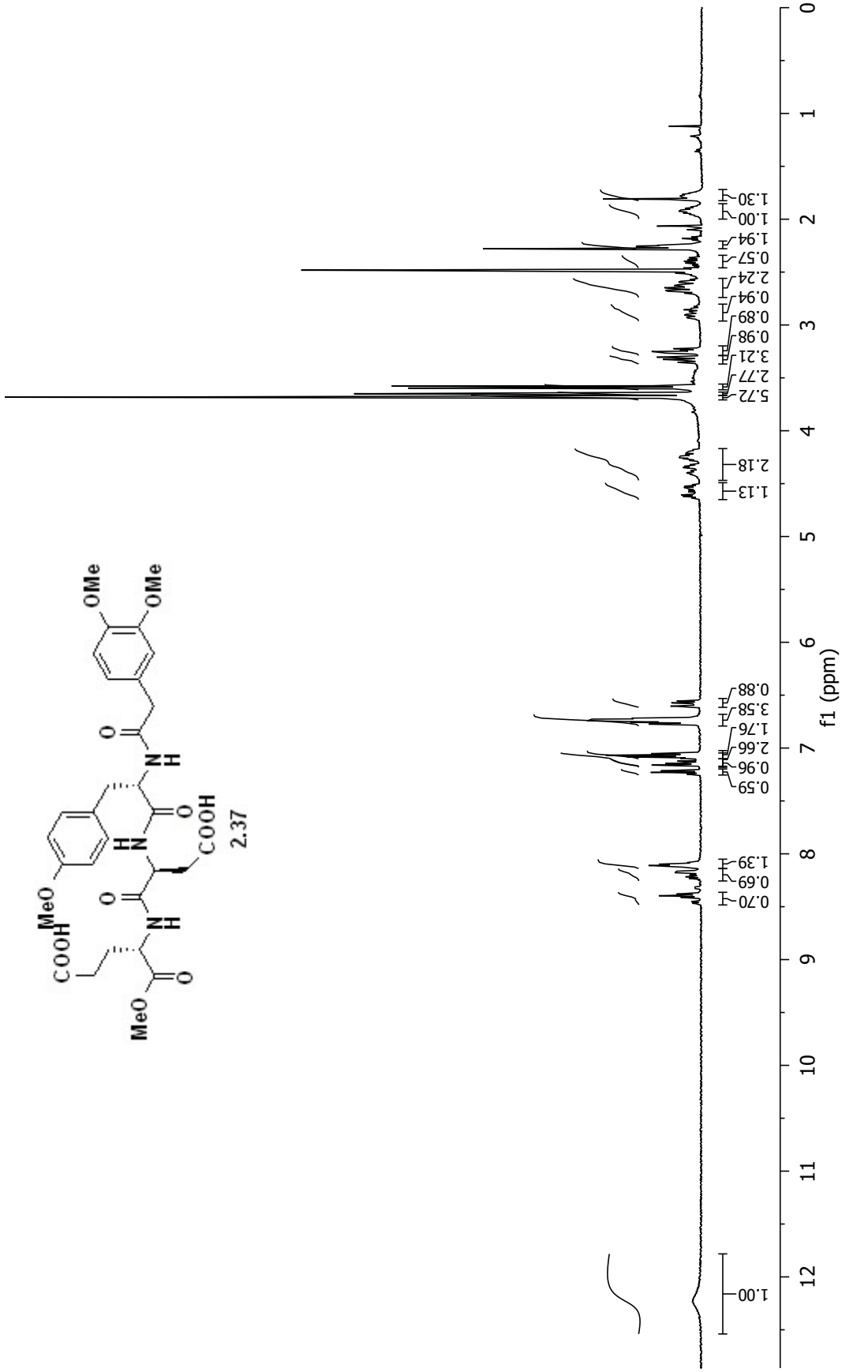

MeO
COOH
O
N
H
O
N
H
O
N
H
O
HOOC
MeO
2.38
f1 (ppm)
0
10
20
30
40
50
60
70
80
90
100
110
120
130
140
150
160
170
180
190

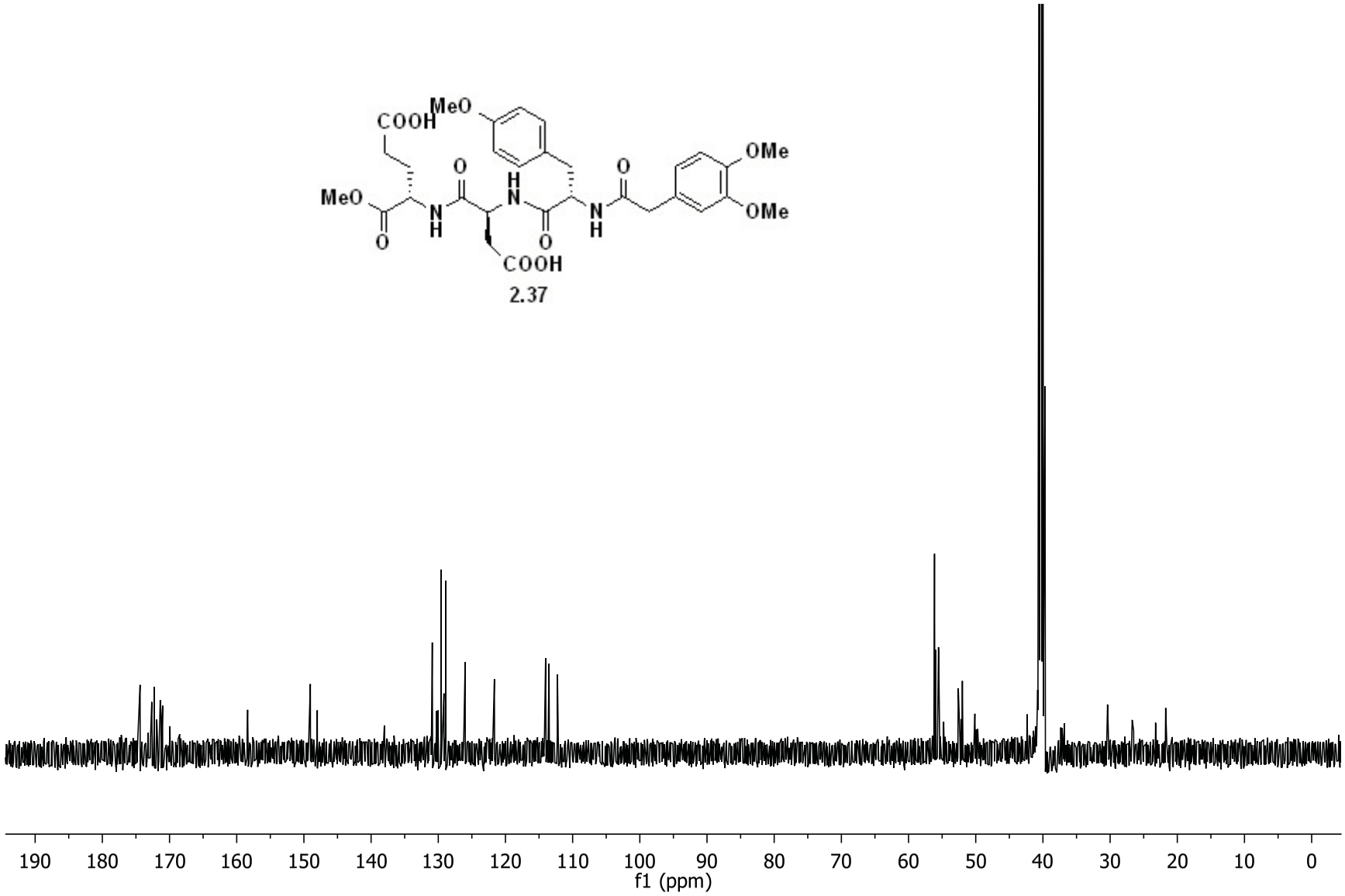

2.40
f1 (ppm)


2.40
f1 (ppm)


COOH
MeO
MeO
O
N
H
O
H
N
O
N
H
O
Cl
Cl
O
COOH
2.36
0.71
1.96
1.04
0.94
0.94
3.02
1.93
0.99
0.95
1.01
2.94
2.99
1.06
1.07
0.96
2.07
0.85
2.43
1.13
0.60
12
11
10
9
8
7
6
5
4
3
2
1
f1 (ppm)

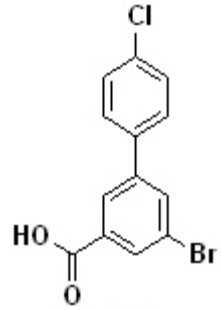

Cl
Cl
O
N
H
O
N
H
O
COOH
2.36
MeO
COOH
O
N
H
O
MeO
f1 (ppm)
0
10
20
30
40
50
60
70
80
90
100
110
120
130
140
150
160
170
180
190


OMe
OMe
O
N
H
O
N
H
O
COOH
2.37
MeO
COOH
O
N
H
O
MeO
1.00
0.70
0.69
1.39
0.59
0.96
2.66
1.76
3.58
0.88
1.13
2.18
5.72
2.77
3.21
0.98
0.89
0.94
2.24
0.57
1.94
1.00
1.30
f1 (ppm)

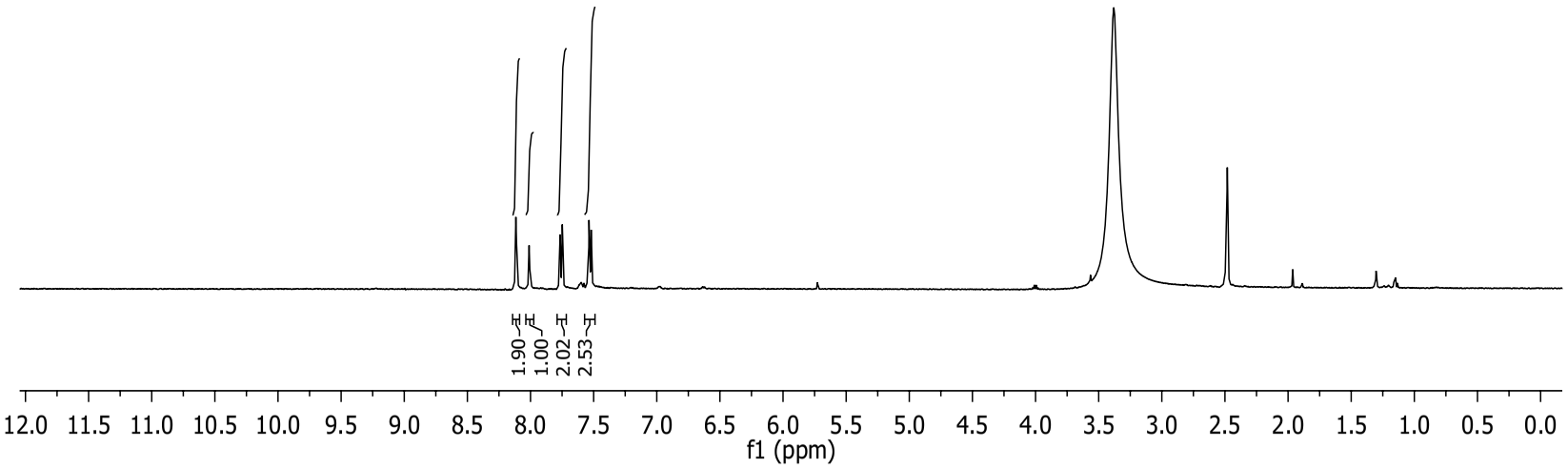

COOH
MeO
OMe
OMe
MeO
O
H
N
N
H
O
O
O
O
N
H
COOH
2.37
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
f1 (ppm)

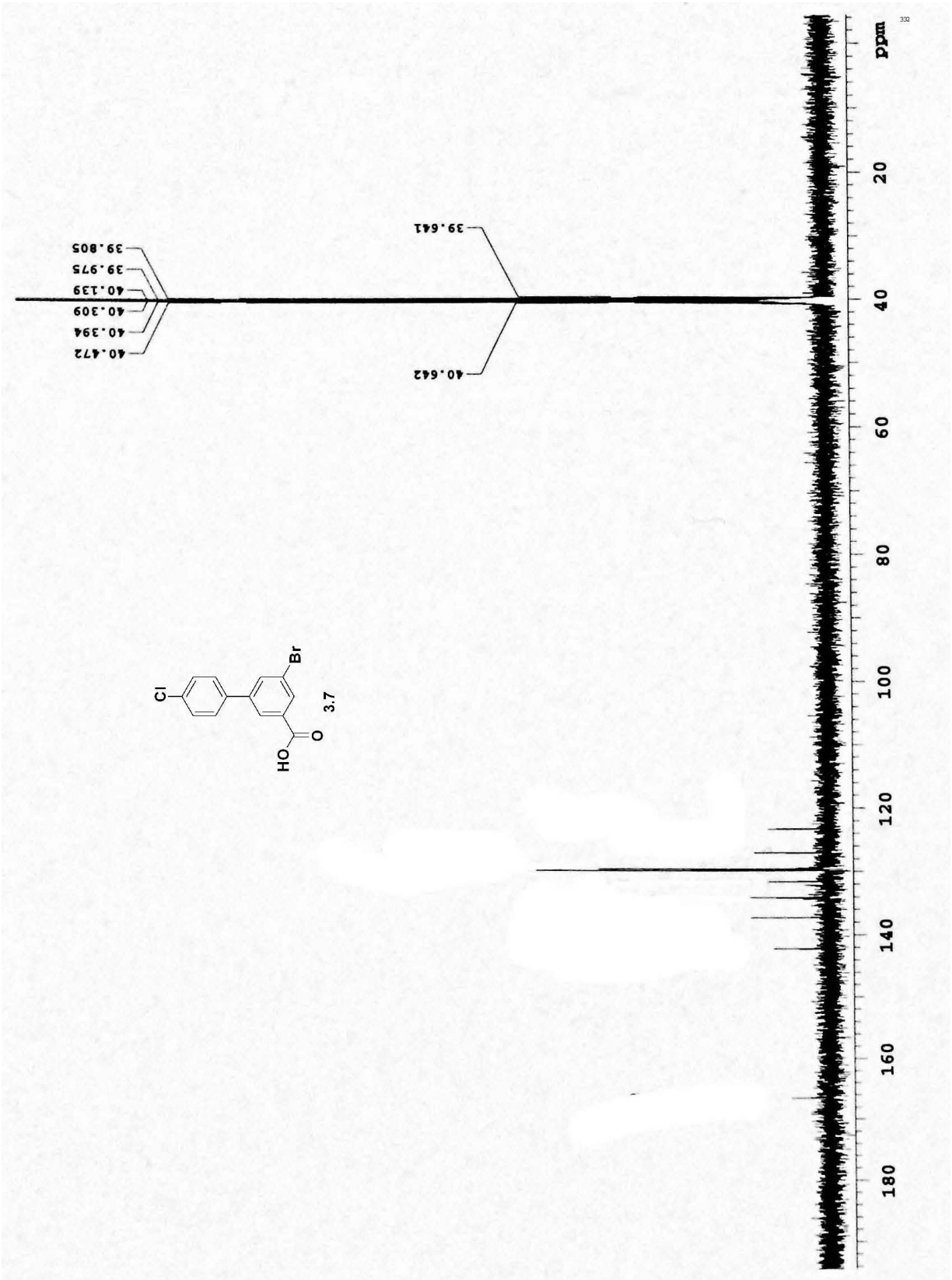
APPENDIX C

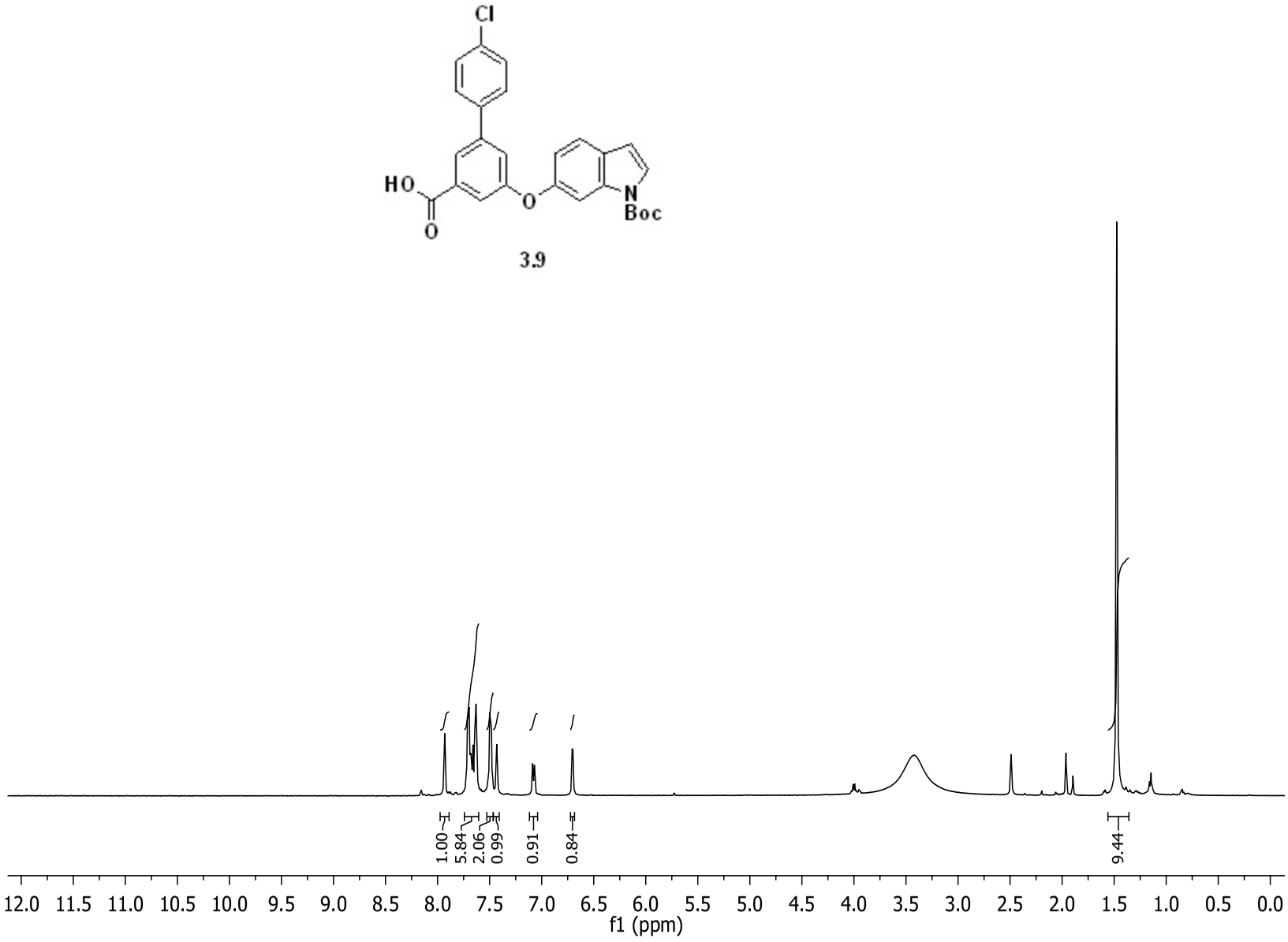
^{1}H AND ^{13}C NMR SPECTRA OF CHAPTER 3

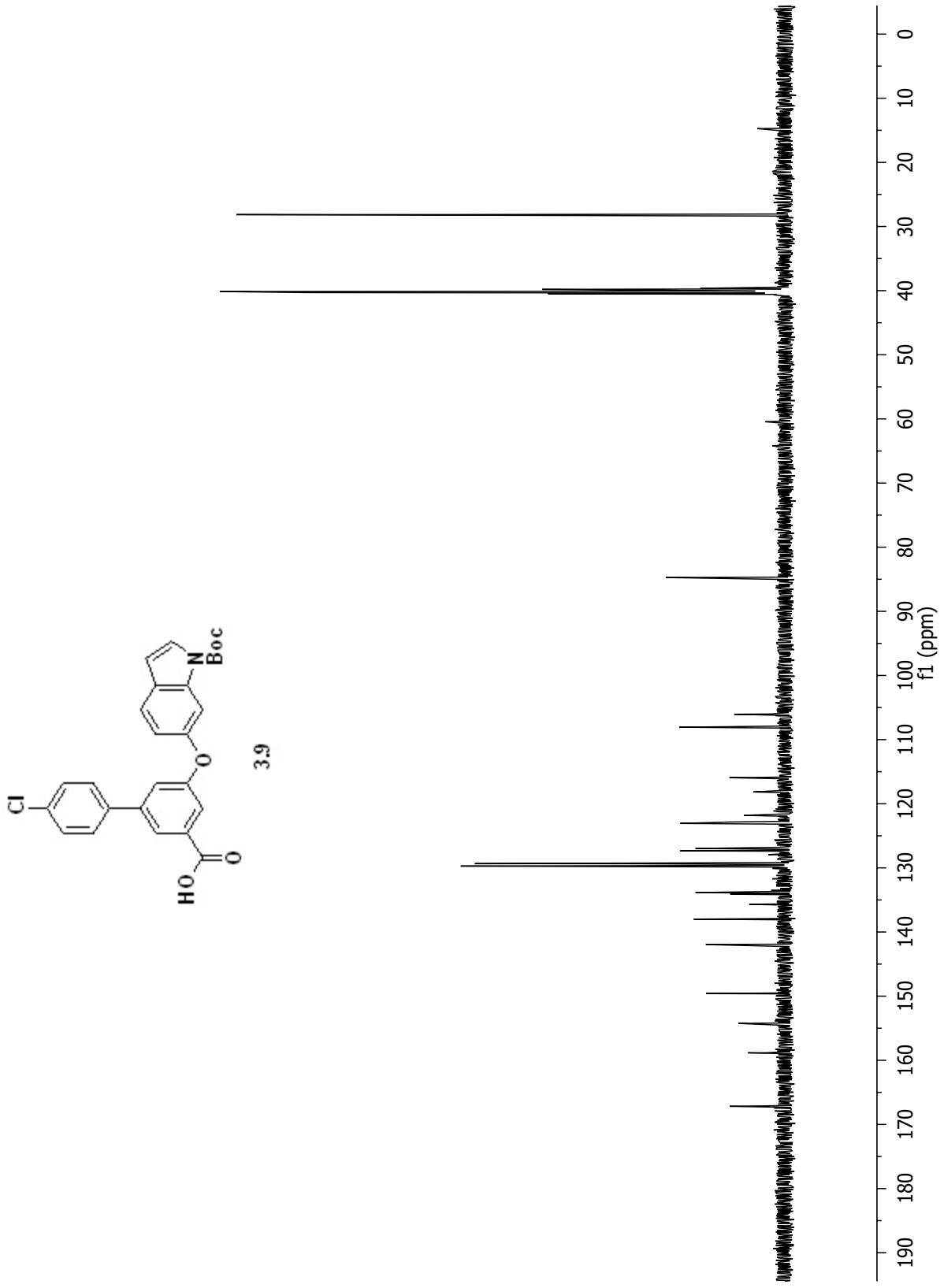

Cl
HO
Br
O

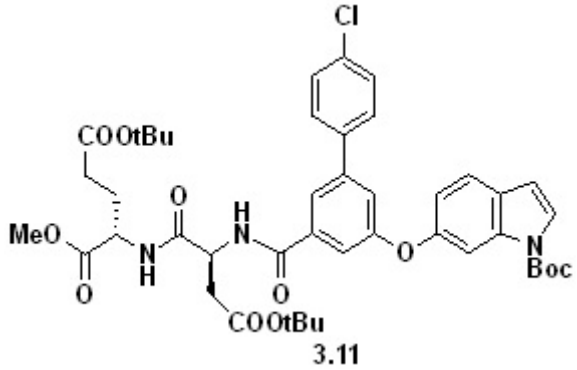
3.7


1.90
1.00
2.02
2.53
12.0 11.5 11.0 10.5 10.0 9.5 9.0 8.5 8.0 7.5 7.0 6.5 6.0 5.5 5.0 4.5 4.0 3.5 3.0 2.5 2.0 1.5 1.0 0.5 0.0
f1 (ppm)

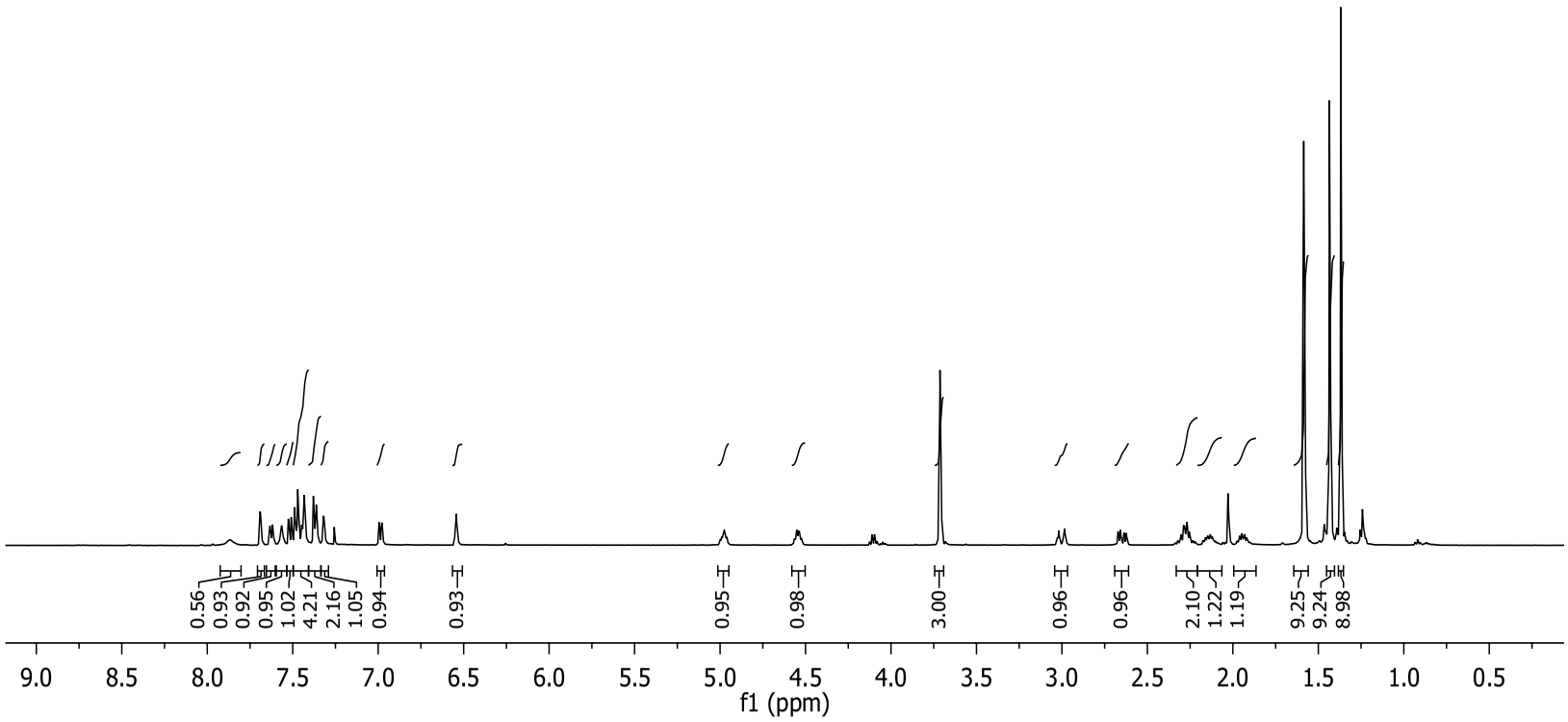

39.805
39.975
40.139
40.309
40.394
40.472
39.641
40.642
Cl
Br
HO
O
3.7
ppm
20
40
60
80
100
120
140
160
180

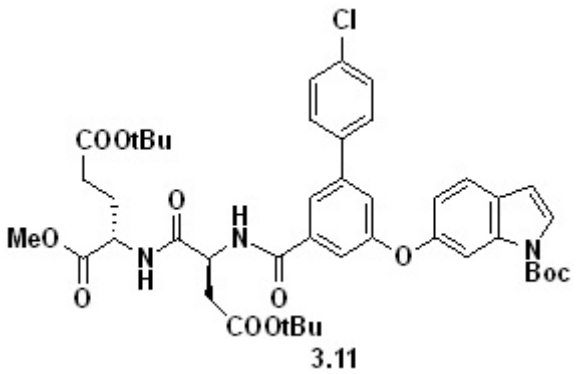

Cl
HO
O
O
N
Boc
3.9
1.00
5.84
2.06
0.99
0.91
0.84
9.44
12.0 11.5 11.0 10.5 10.0 9.5 9.0 8.5 8.0 7.5 7.0 6.5 6.0 5.5 5.0 4.5 4.0 3.5 3.0 2.5 2.0 1.5 1.0 0.5 0.0
f1 (ppm)


Cl
O
HO
O
N
Boc
3.9
0
10
20
30
40
50
60
70
80
90
100
110
120
130
140
150
160
170
180
190
f1 (ppm)

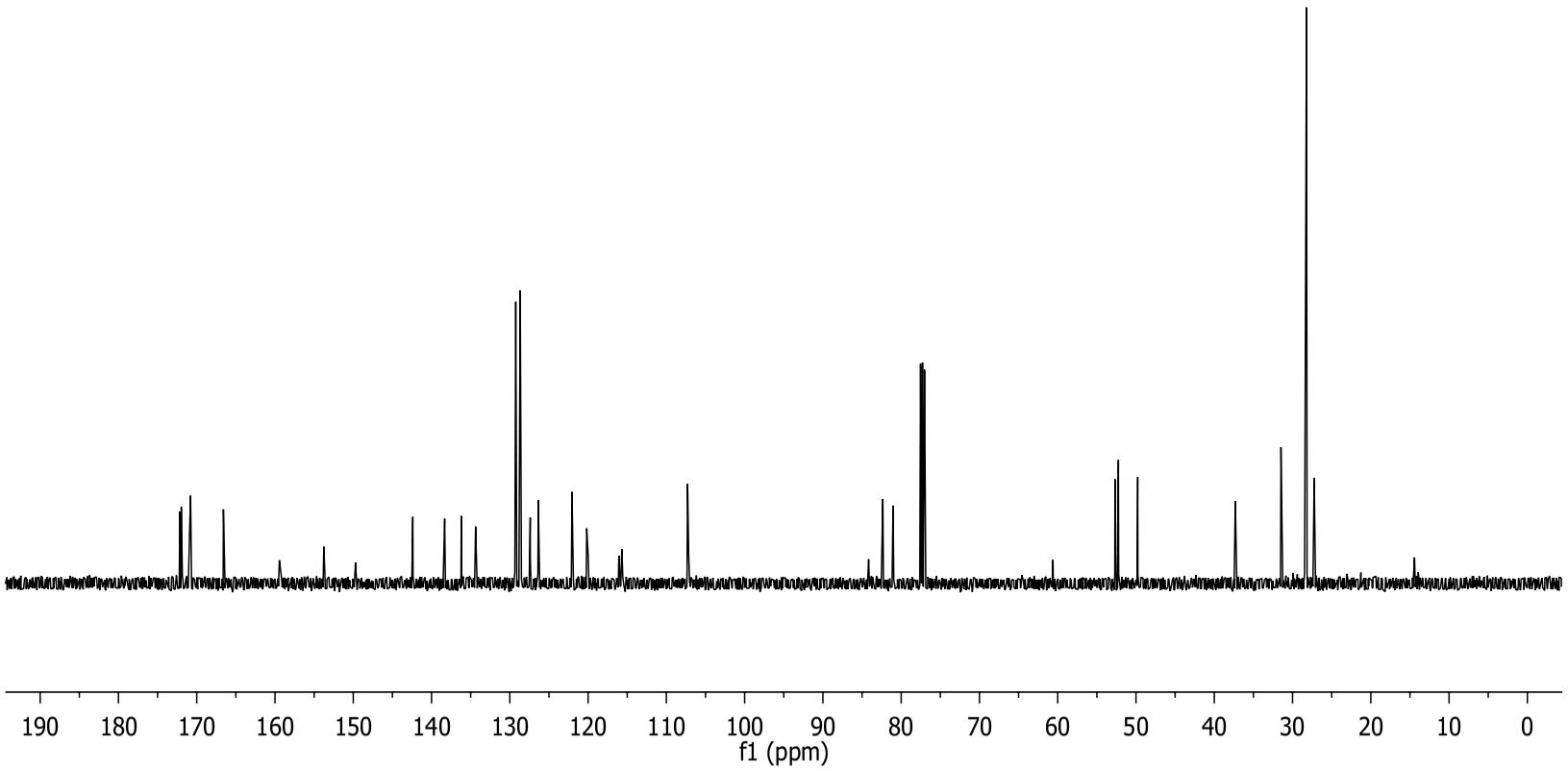

Cl
COOtBu
MeO
O
N
H
O
H
N
O
O
N
Boc
O
COOtBu

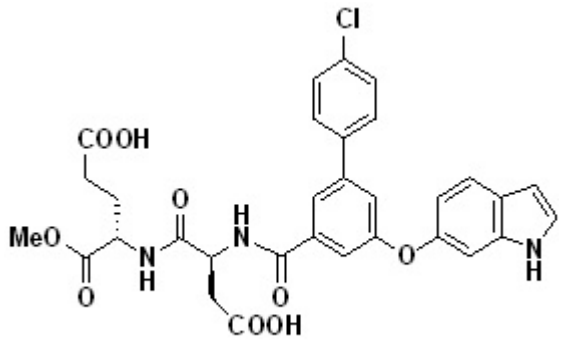
3.11


0.56
0.93
0.92
0.95
1.02
4.21
2.16
1.05
0.94
0.93
0.95
0.98
3.00
0.96
0.96
2.10
1.22
1.19
9.25
9.24
8.98
9.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
f1 (ppm)

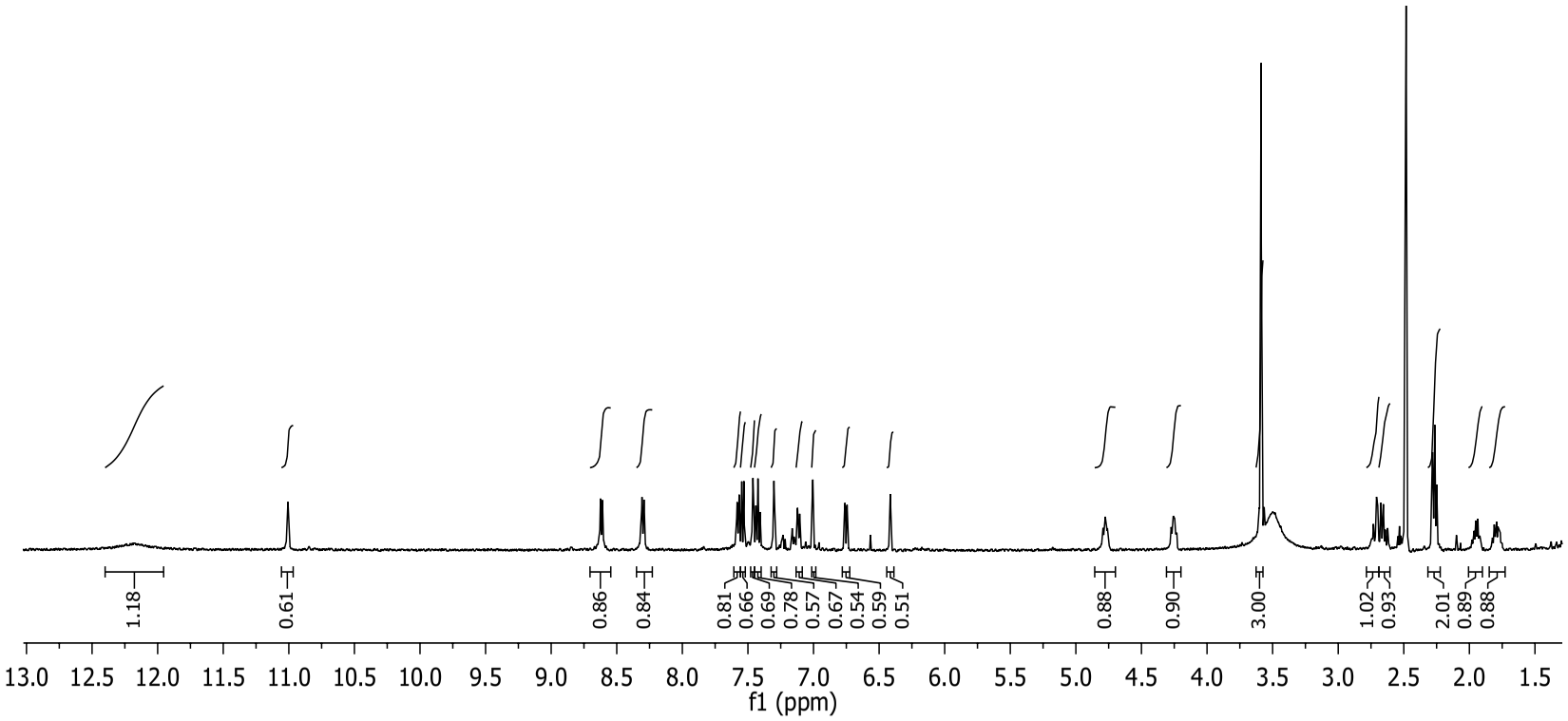

Cl
COOtBu
MeO
O
H
N
N
H
O
O
O
N
Boc
COOtBu

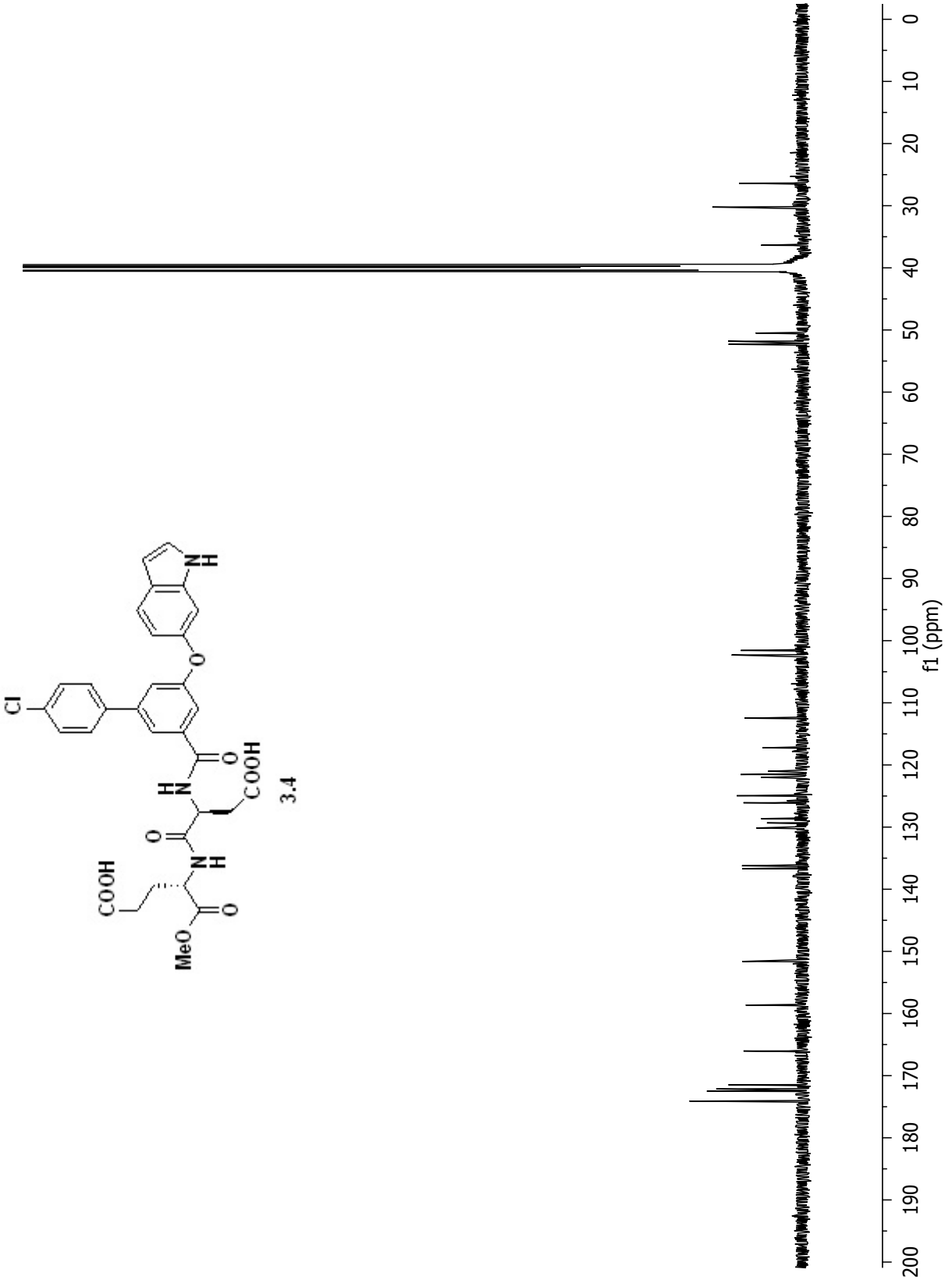
3.11

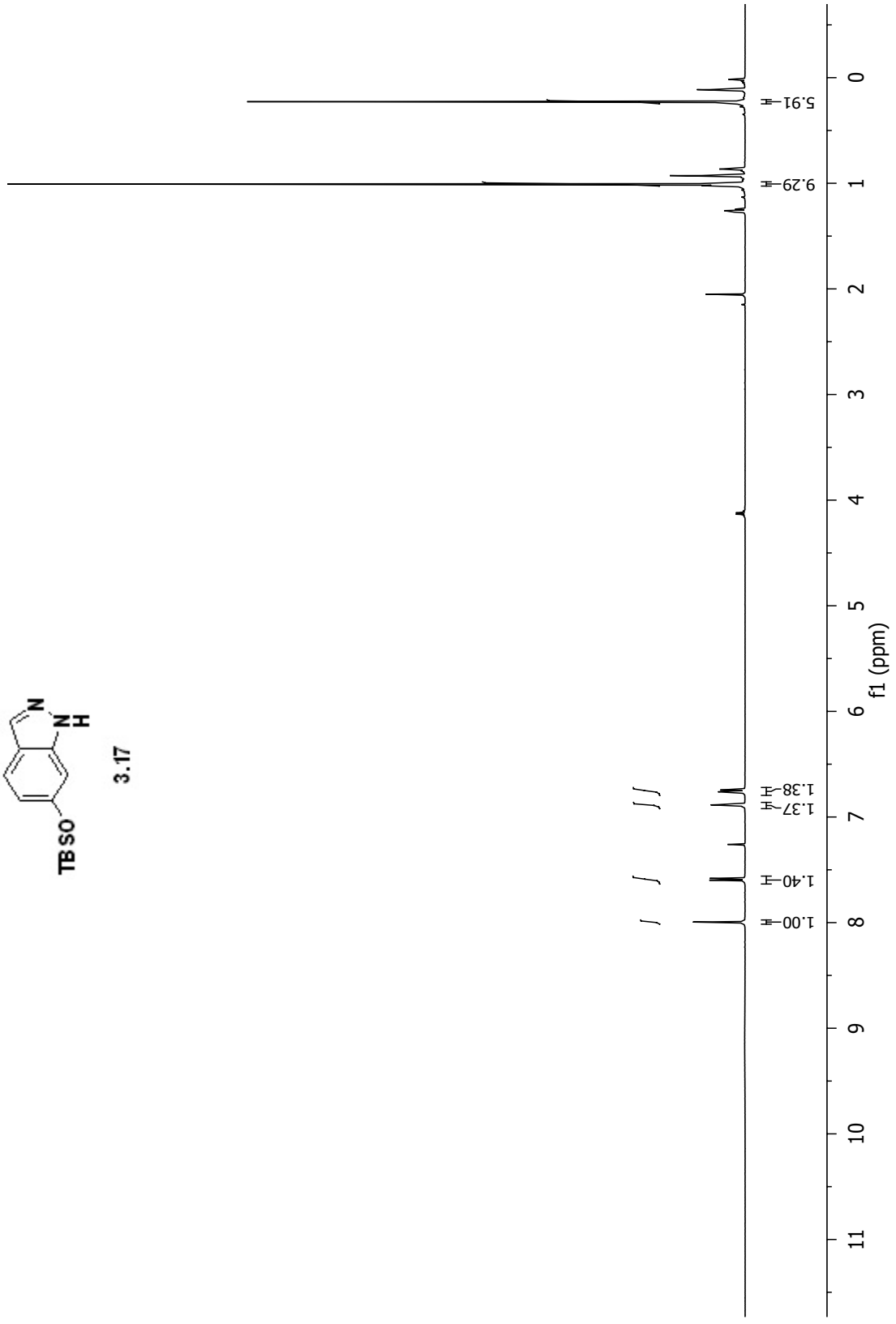

190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
f1 (ppm)

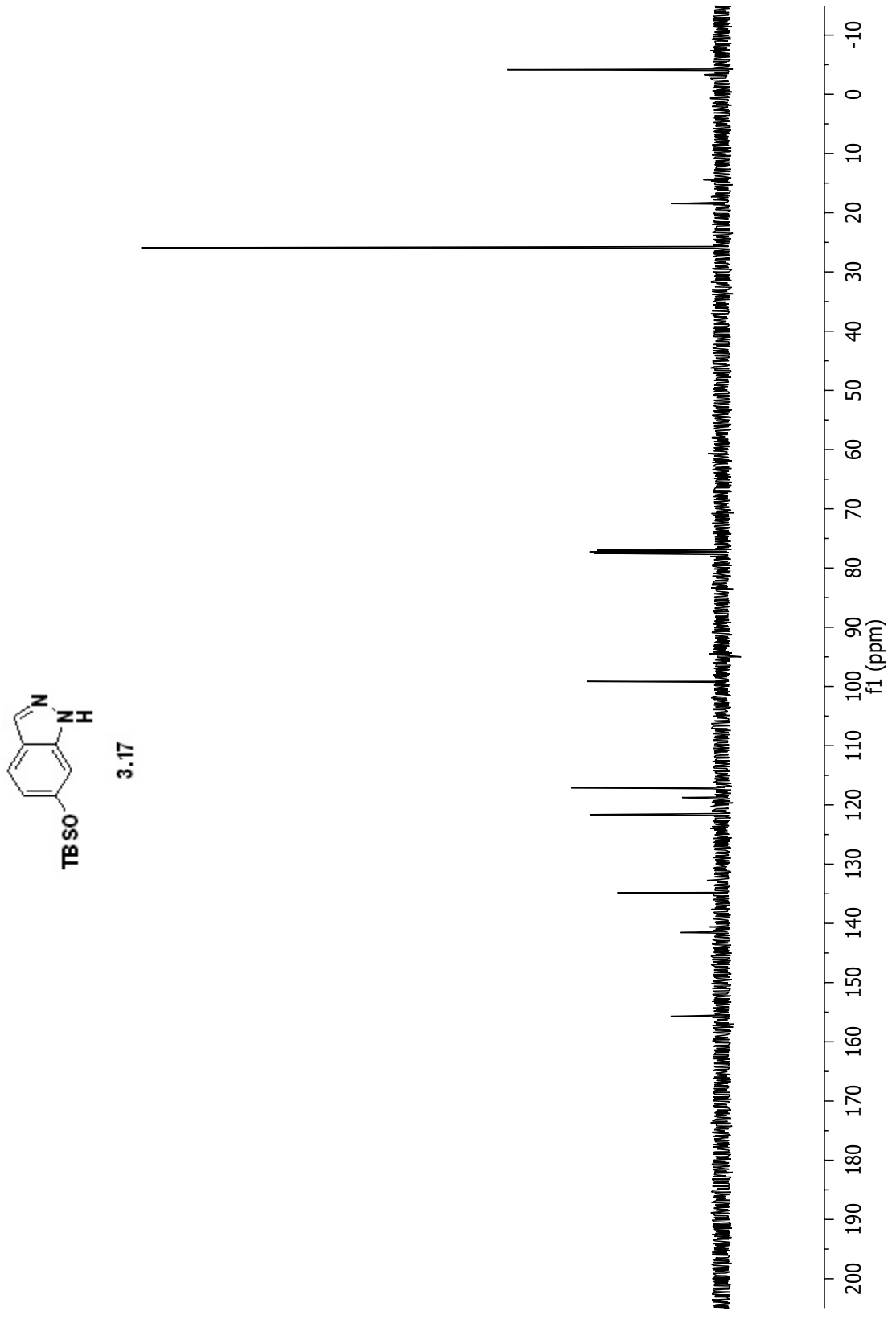

Cl
COOH
MeO
O
N
H
O
H
N
O
O
N
H
COOH

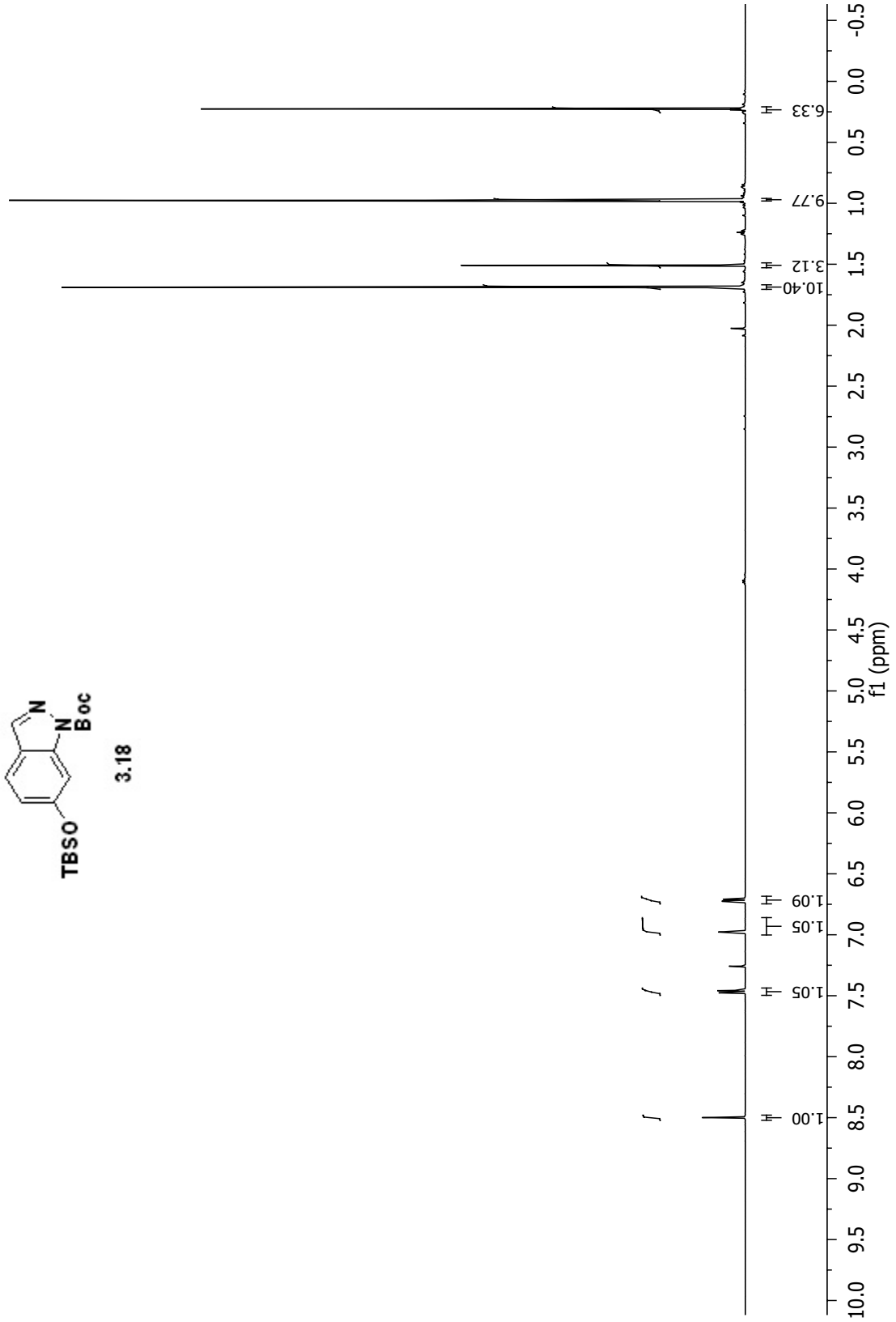
3.4

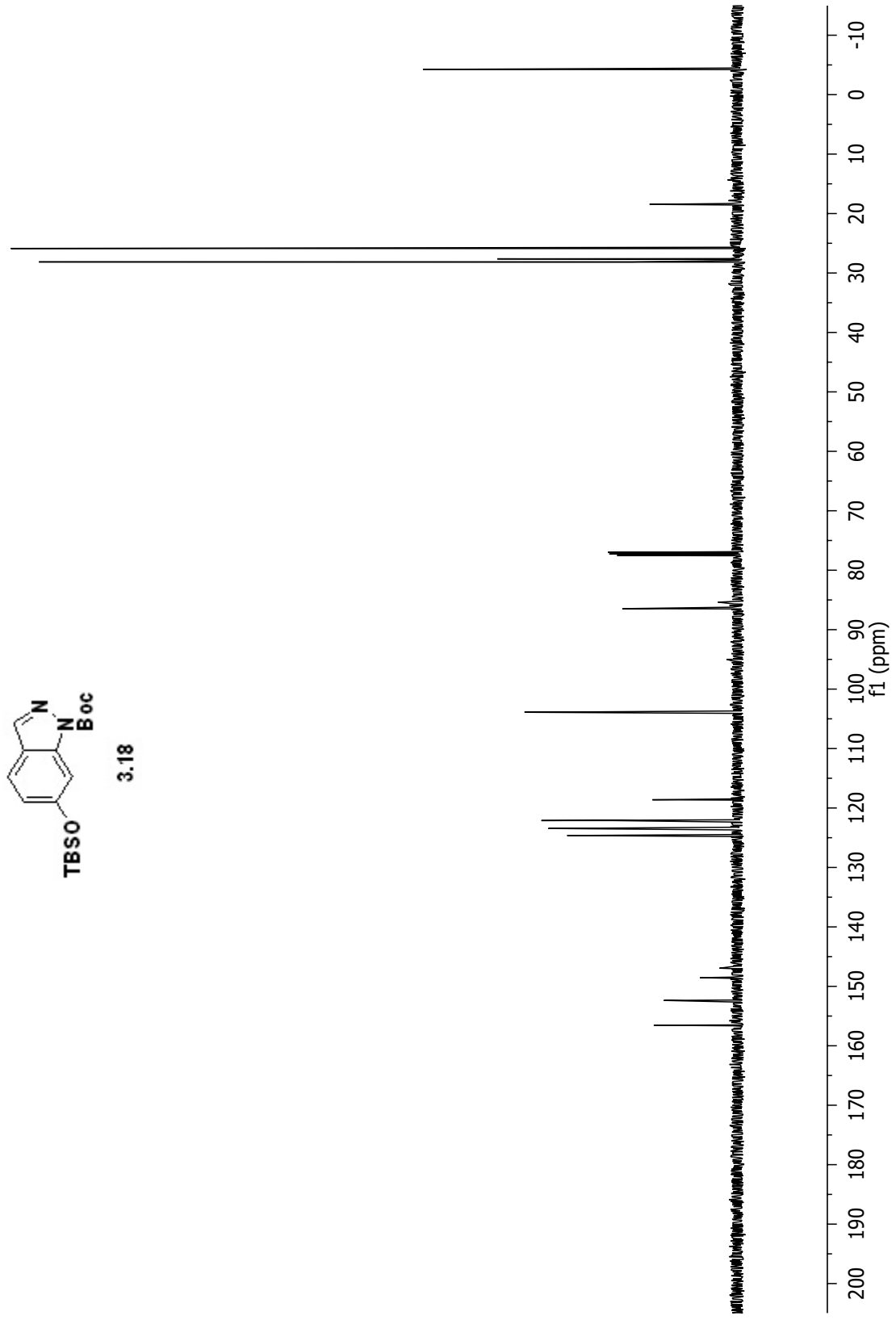

1.18
0.61
0.86
0.84
0.81
0.66
0.69
0.78
0.57
0.67
0.54
0.59
0.51
0.88
0.90
3.00
1.02
0.93
2.01
0.89
0.88
13.0 12.5 12.0 11.5 11.0 10.5 10.0 9.5 9.0 8.5 8.0 7.5 7.0 6.5 6.0 5.5 5.0 4.5 4.0 3.5 3.0 2.5 2.0 1.5
f1 (ppm)

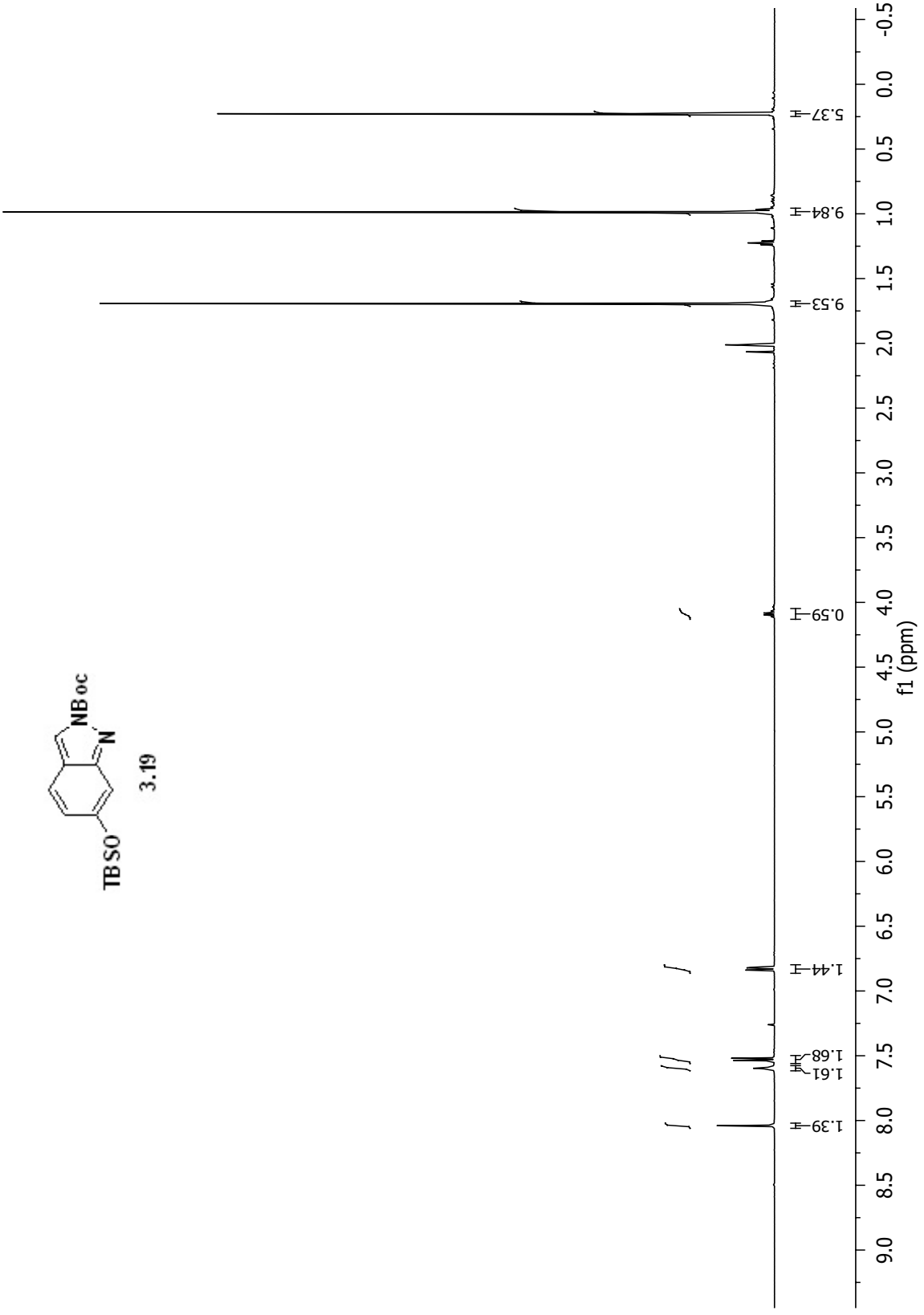

Cl
O
N
H
O
N
H
COOH
O
O
N
H
COOH
O
MeO
3.4
f1 (ppm)
0
10
20
30
40
50
60
70
80
90
100
110
120
130
140
150
160
170
180
190
200

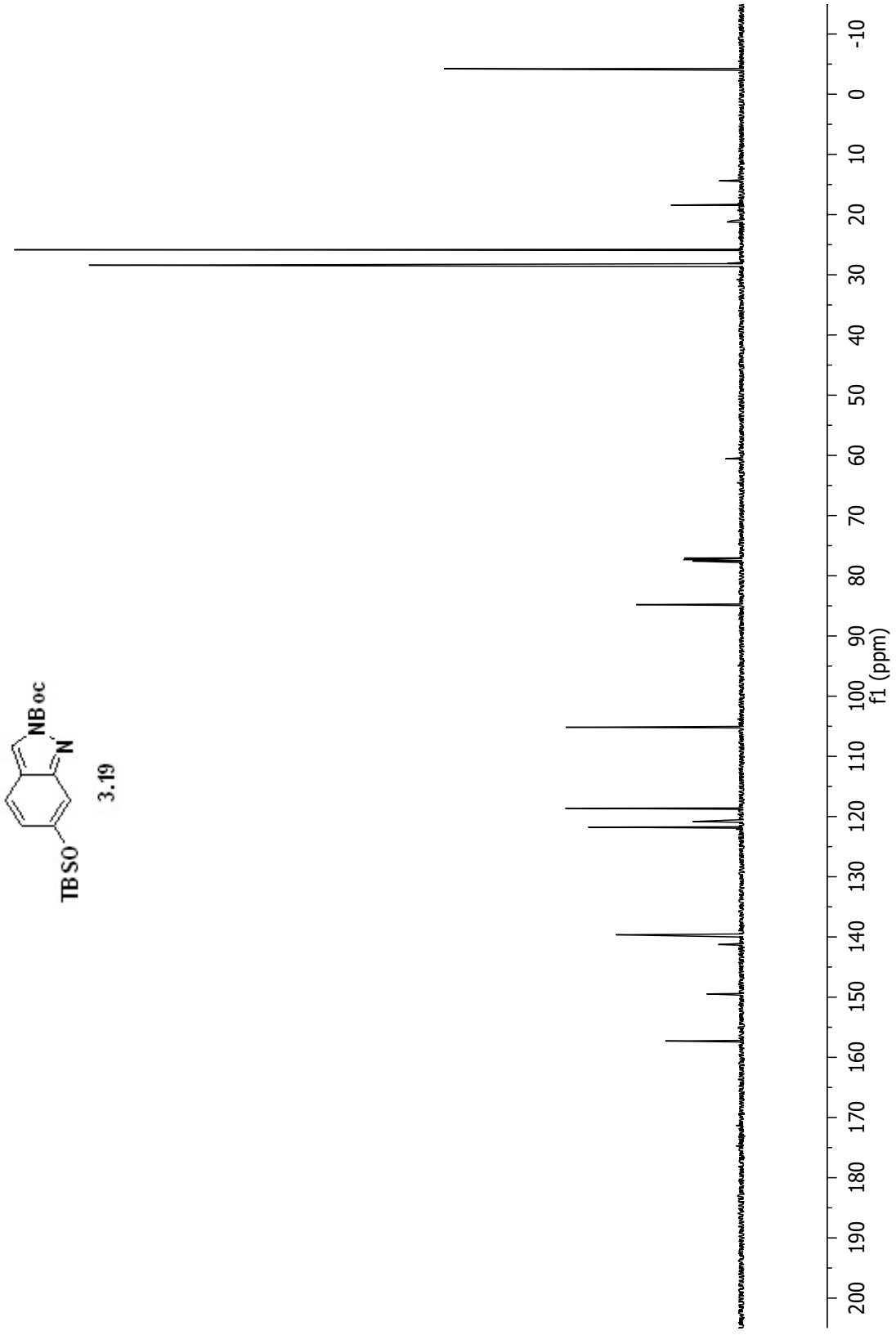

TBSO
N
N
H
3.17
5.91
9.29
1.00
1.40
1.37
1.38
f1 (ppm)

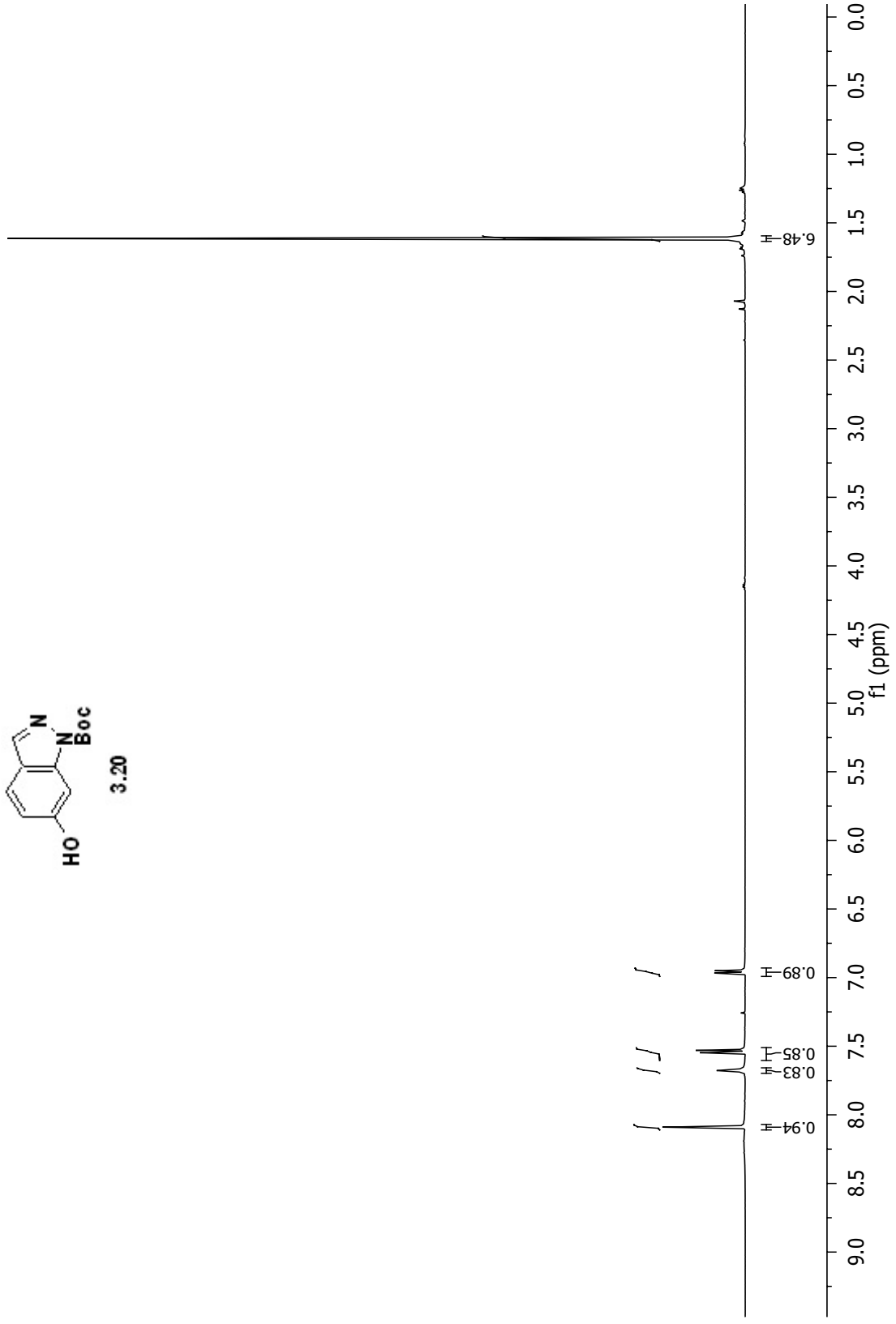

TBSO
N
N
H
3.17
f1 (ppm)
-10
0
10
20
30
40
50
60
70
80
90
100
110
120
130
140
150
160
170
180
190
200

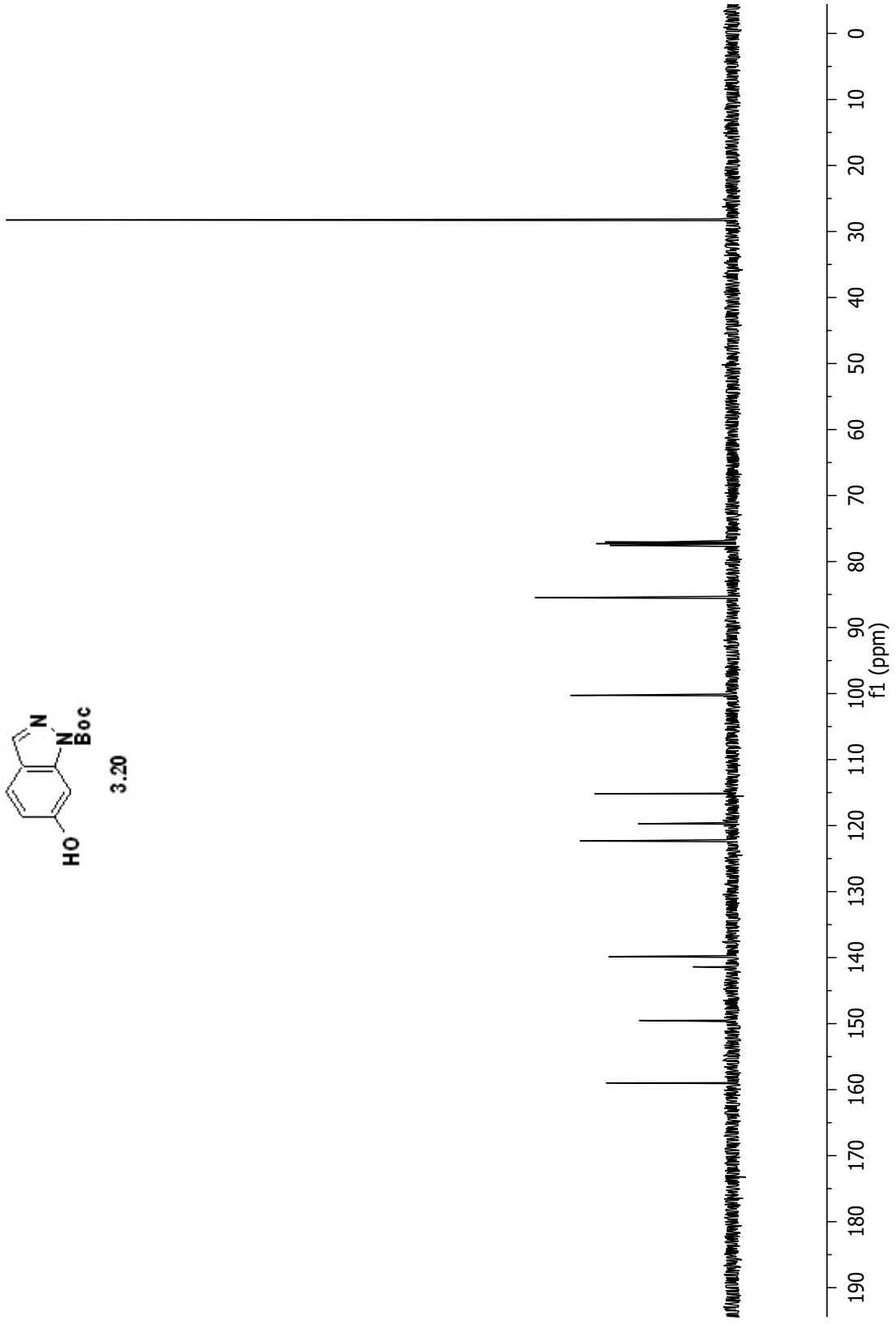

TBSO
N
N
Boc
3.18
1.00
1.05
1.05
1.09
10.40
3.12
9.77
6.33
10.0
9.5
9.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
-0.5
f1 (ppm)

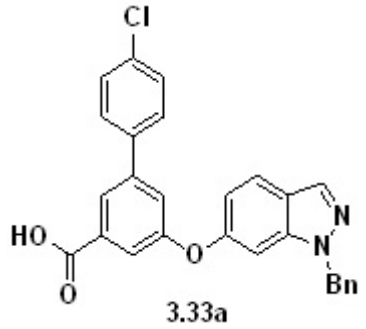

N
N
Boc
TBSO
3.18
200
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
-10
f1 (ppm)


NBoc
N
TBSO
3.19
9.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
-0.5
f1 (ppm)
1.39
1.61
1.68
1.44
0.59
9.53
9.84
5.37

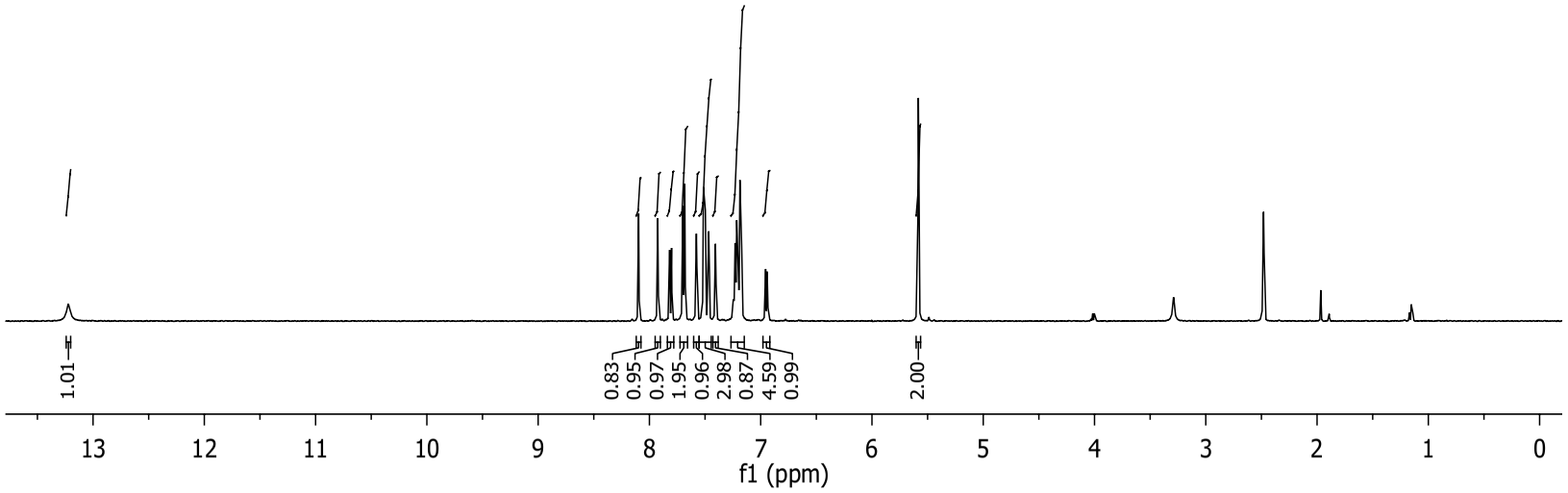

NBoc
N
TBSO
3.19
200
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
-10
f1 (ppm)

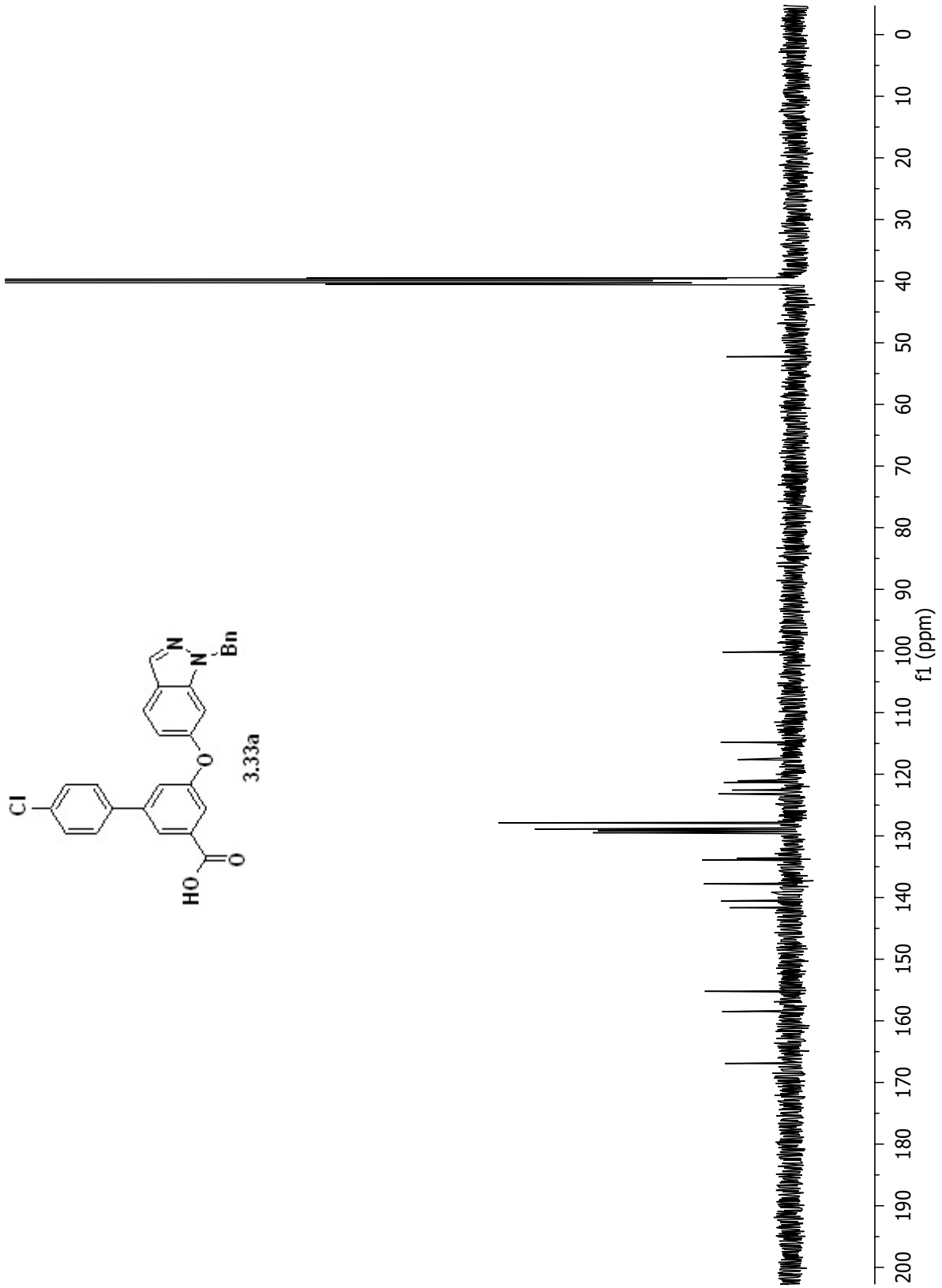

N
N
Boc
HO
3.20
6.48
0.89
0.85
0.83
0.94
f1 (ppm)

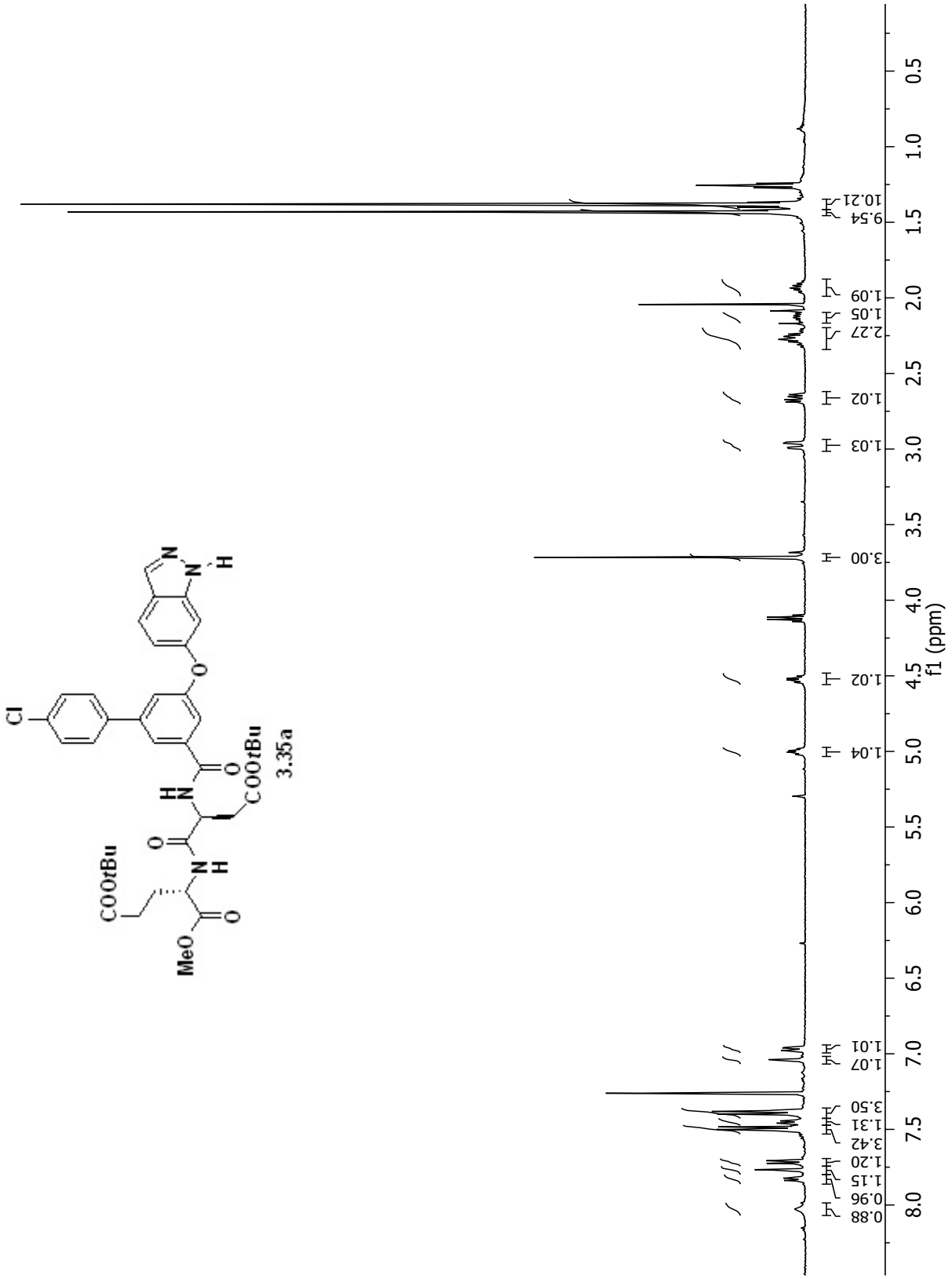

N
N
Boc
HO
3.20
0
10
20
30
40
50
60
70
80
90
100
110
120
130
140
150
160
170
180
190
f1 (ppm)

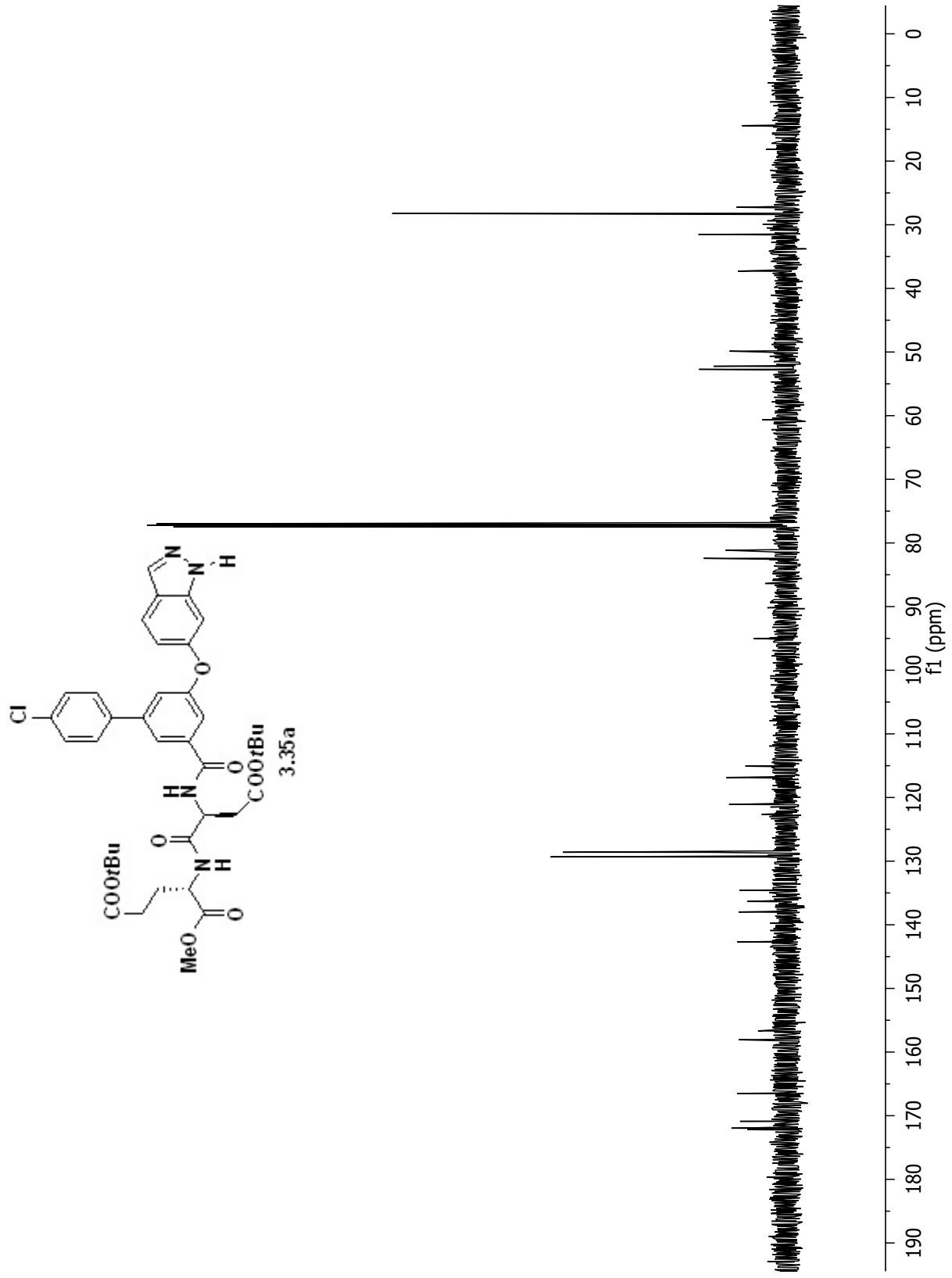

Cl
HO
O
O
N
N
Bn

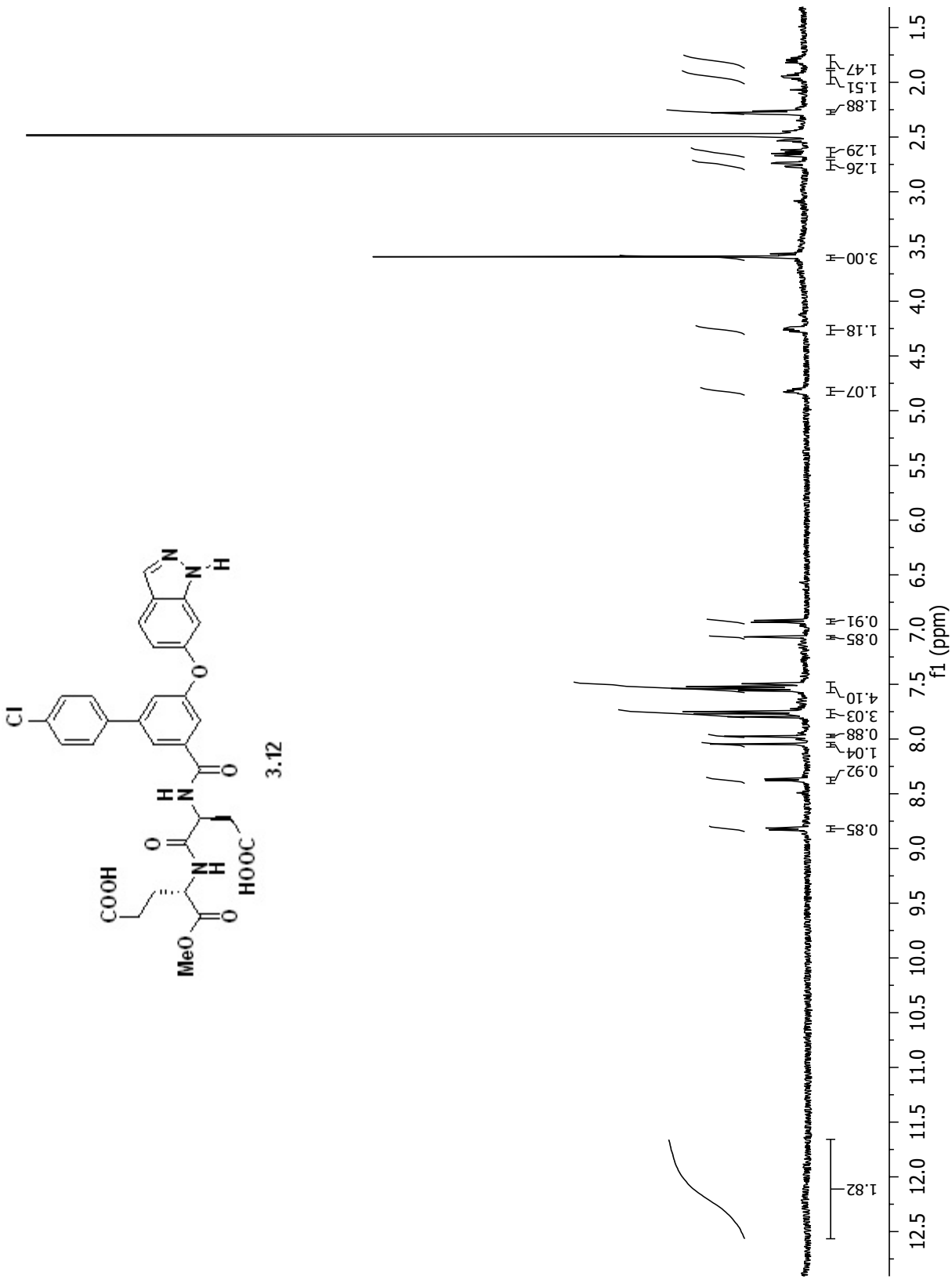
3.33a

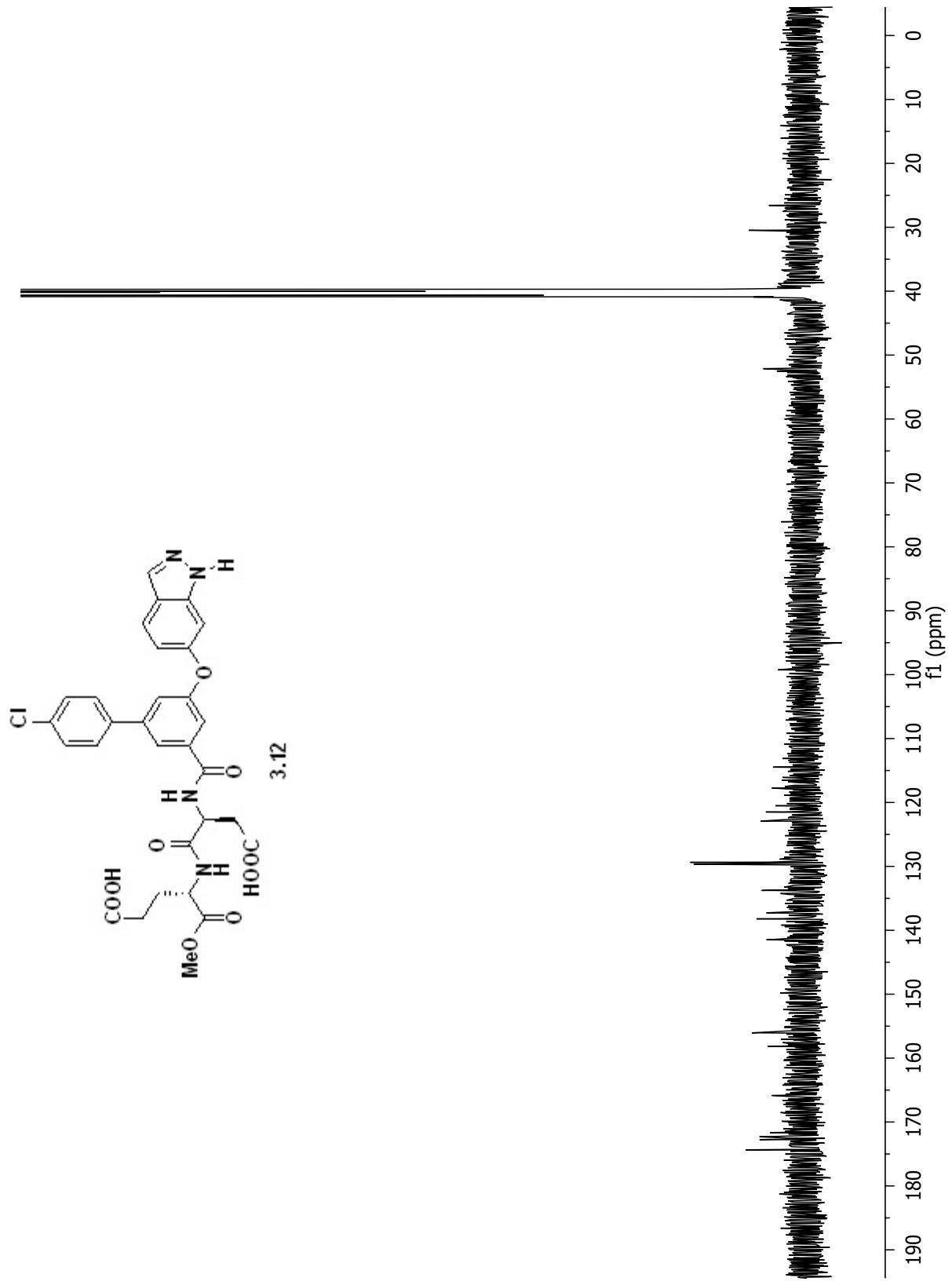

1.01
0.83
0.95
0.97
1.95
0.96
2.98
0.87
4.59
0.99
2.00
13
12
11
10
9
8
7
6
5
4
3
2
1
0
f1 (ppm)

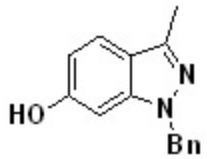

Cl
O
HO
O
N
N
Bn
3.33a
0
10
20
30
40
50
60
70
80
90
100
110
120
130
140
150
160
170
180
190
200
f1 (ppm)


Cl
COOtBu
COOtBu
3.35a
MeO
0.88
0.96
1.15
1.20
3.42
1.31
3.50
1.07
1.01
1.04
1.02
3.00
1.03
1.02
2.27
1.05
1.09
9.54
10.21
f1 (ppm)

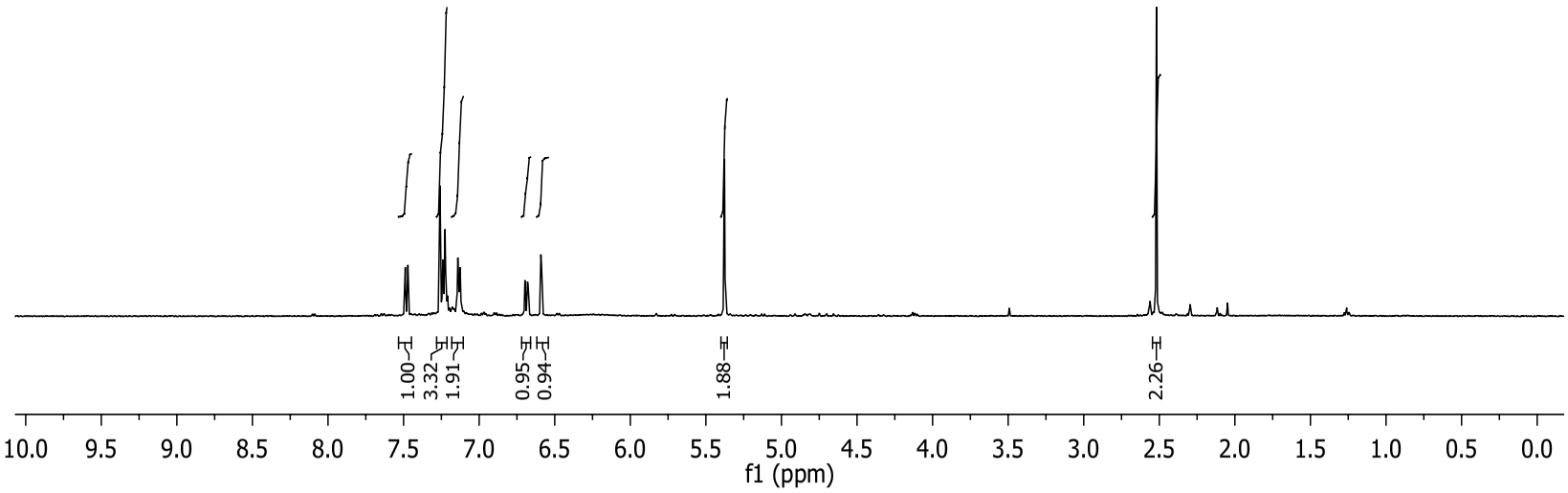

Cl
O
N
H
N
H
O
O
COOtBu
3.35a
N
H
O
COOtBu
N
H
O
MeO
O
0
10
20
30
40
50
60
70
80
90
100
110
120
130
140
150
160
170
180
190
f1 (ppm)

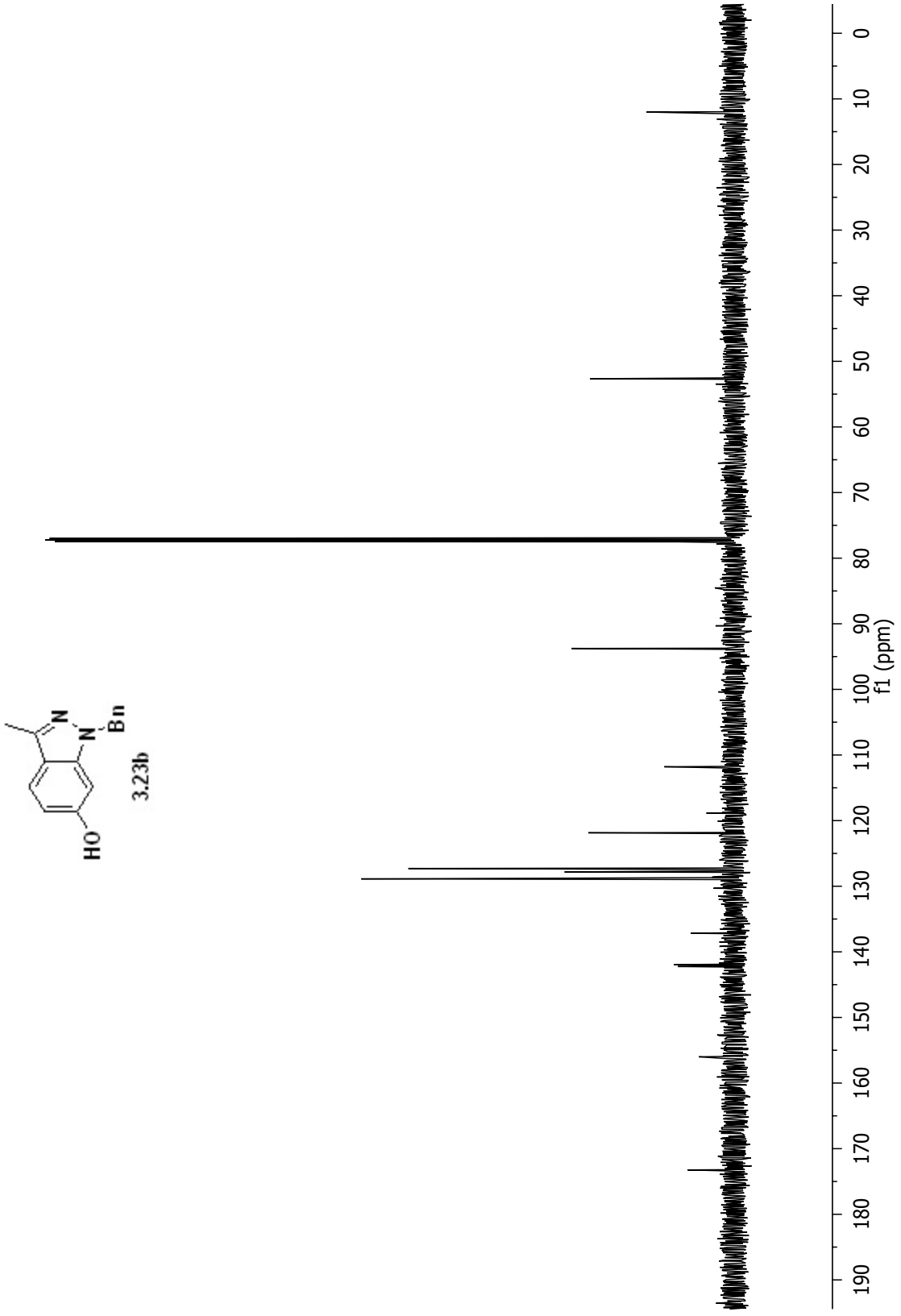

Cl
O
O
H
N
O
N
H
O
COOH
HOOC
MeO
O
O
N
H
N
H
3.12
f1 (ppm)
12.5
12.0
11.5
11.0
10.5
10.0
9.5
9.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.82
0.85
0.92
1.04
0.88
3.03
4.10
0.85
0.91
1.07
1.18
3.00
1.26
1.29
1.88
1.51
1.47

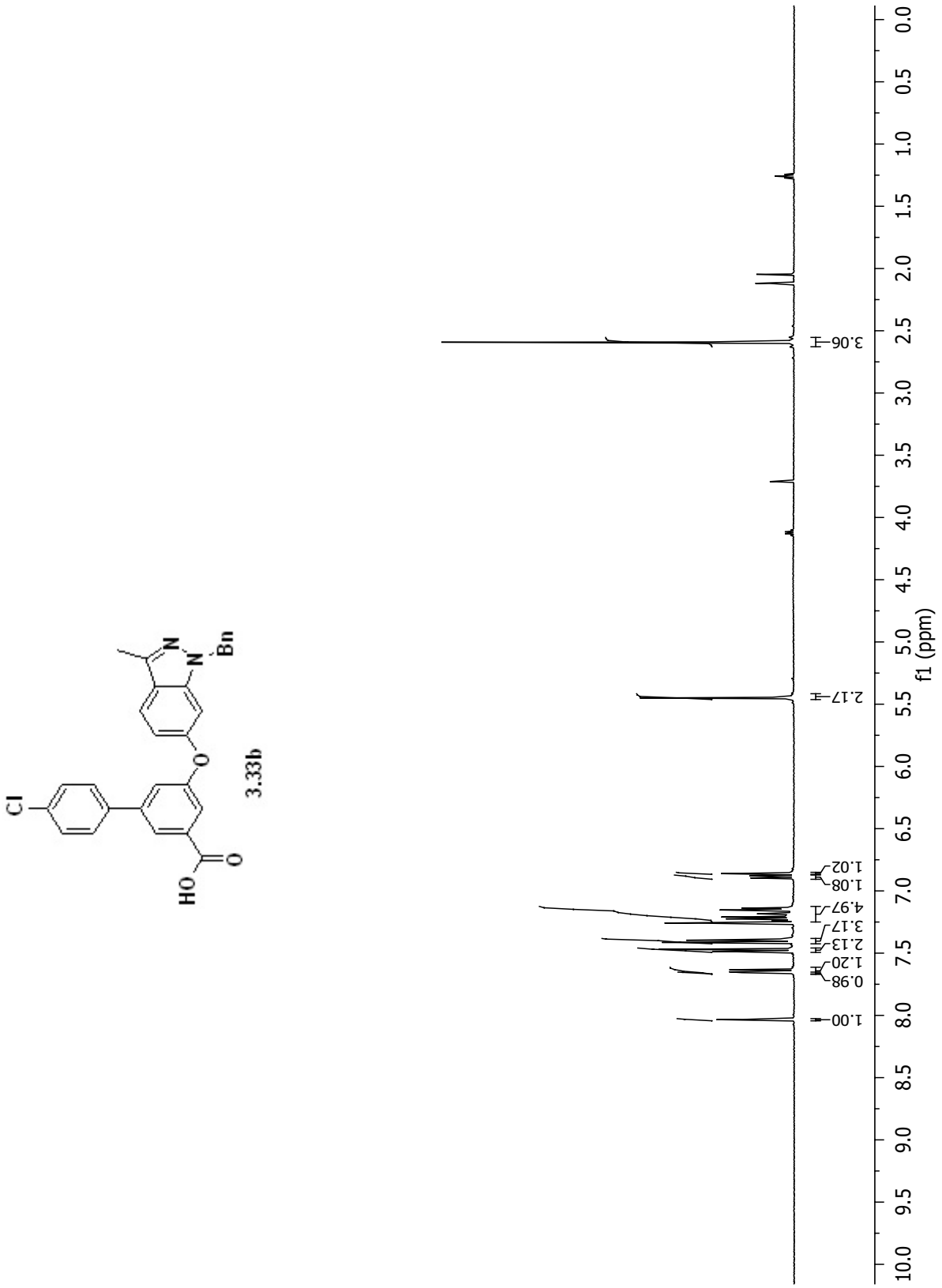

Cl
O
N
H
N
H
O
O
N
H
O
COOH
HOOC
MeO
O
3.12
0
10
20
30
40
50
60
70
80
90
100
110
120
130
140
150
160
170
180
190
f1 (ppm)

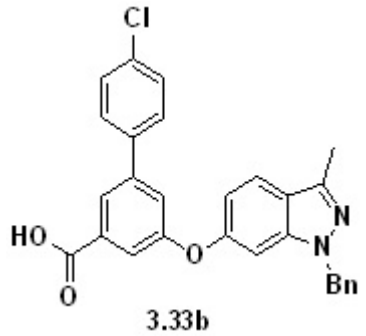

HO
N
N
Bn

3.23b

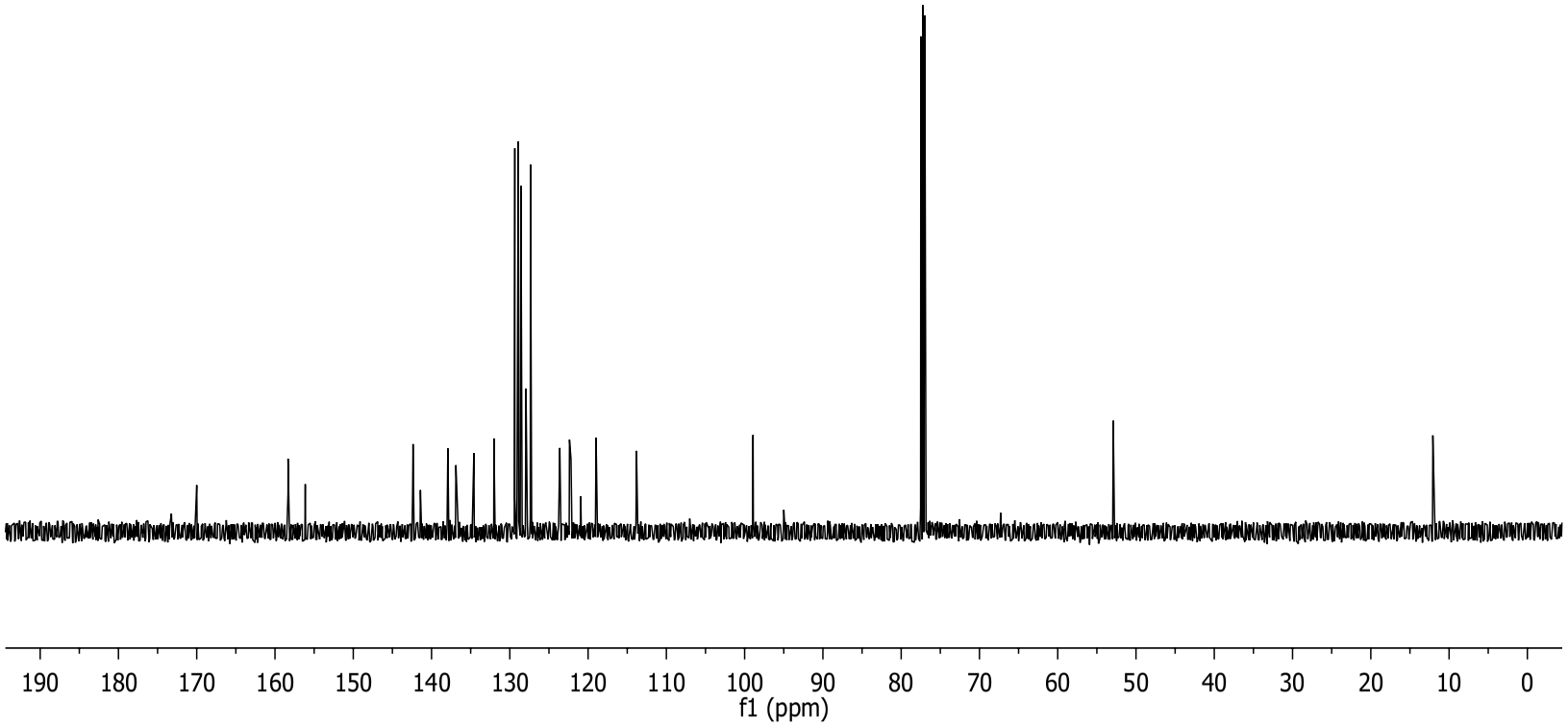

1.00
3.32
1.91
0.95
0.94
1.88
2.26
10.0
9.5
9.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
f1 (ppm)

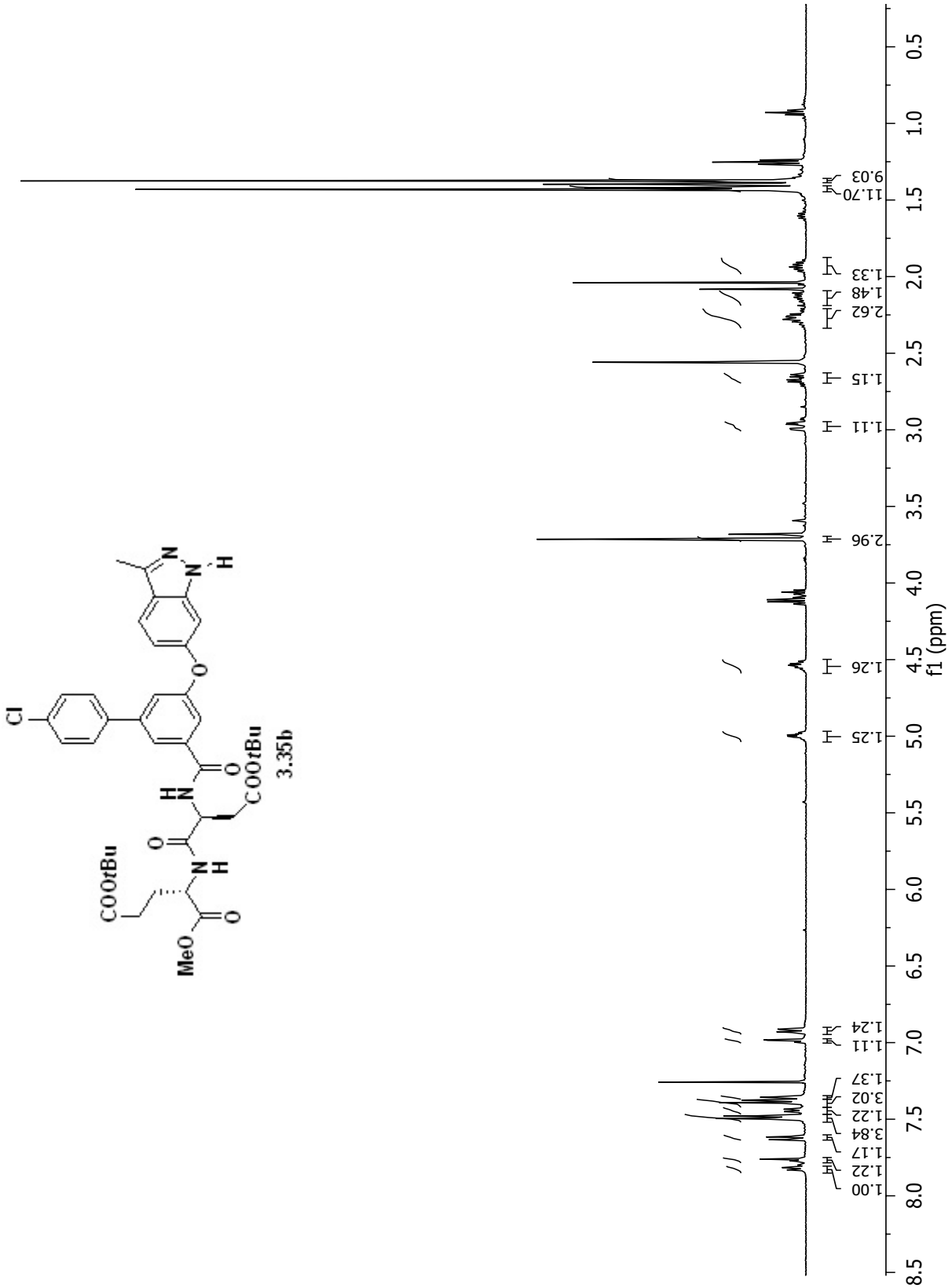

N
N
Bn
HO
3.23b
0
10
20
30
40
50
60
70
80
90
100
110
120
130
140
150
160
170
180
190
f1 (ppm)

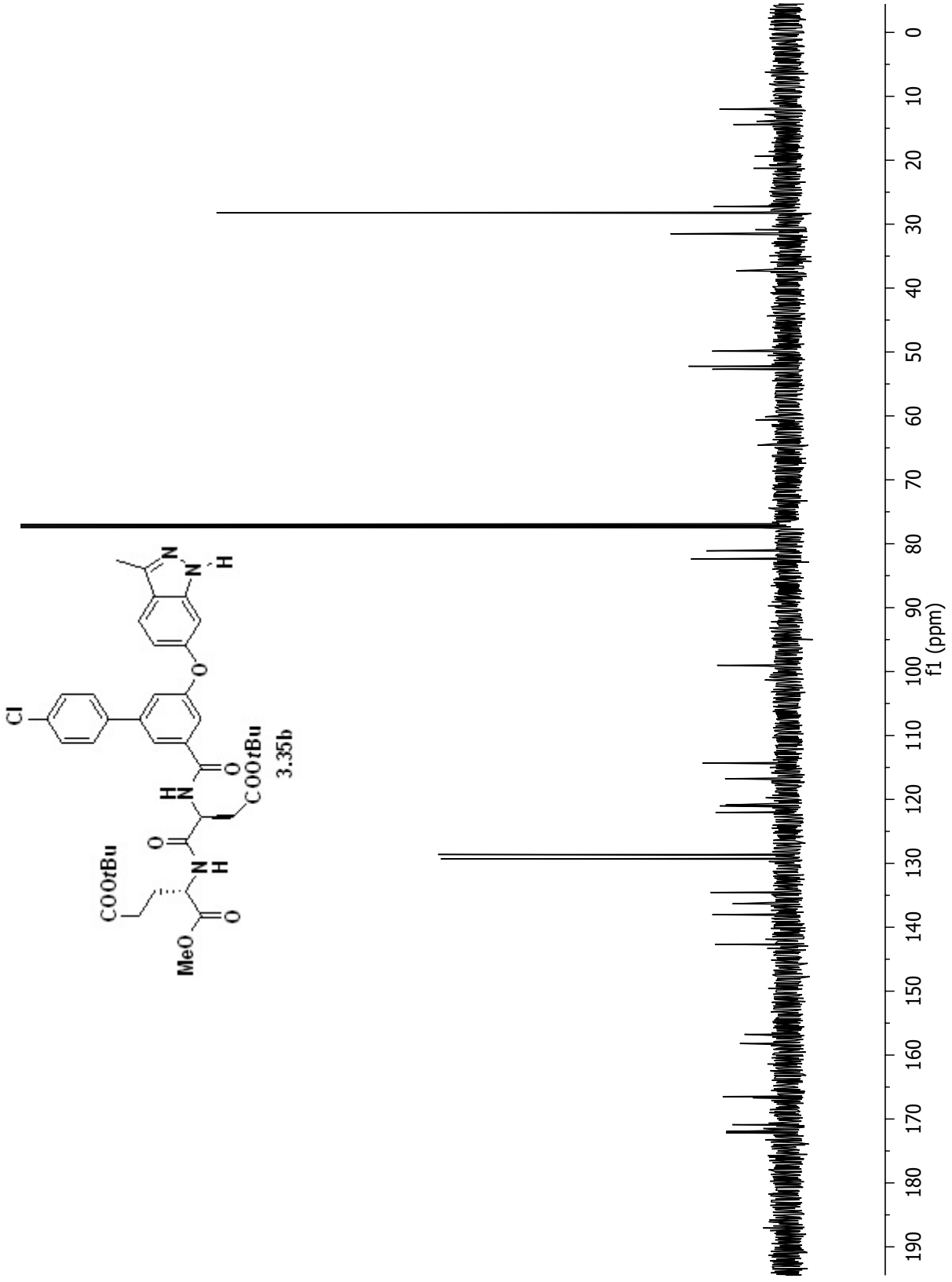

Cl
O
HO
O
N
N
Bn
3.33b
1.00
0.98
1.20
2.13
3.17
4.97
1.08
1.02
2.17
3.06
10.0
9.5
9.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
f1 (ppm)

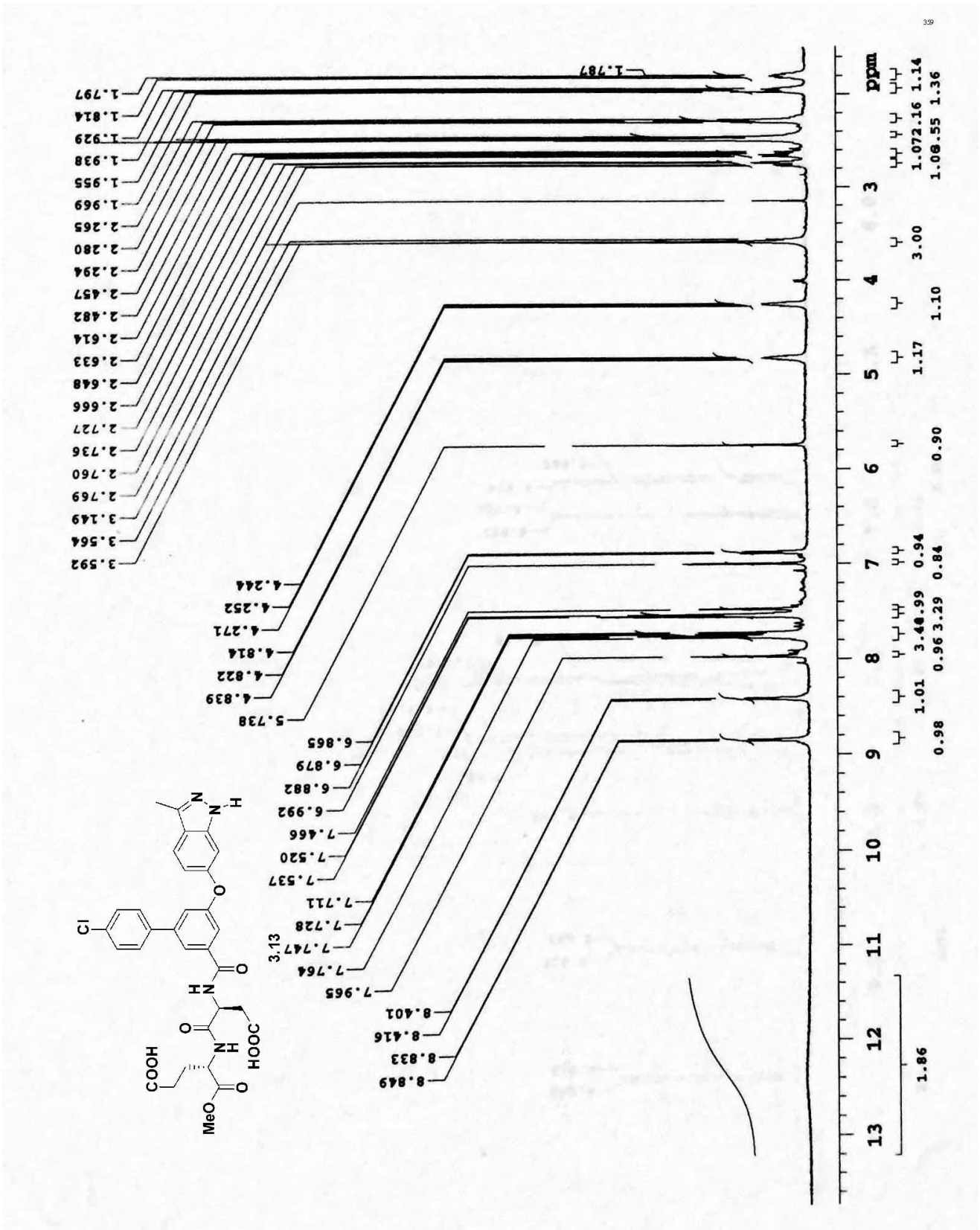

Cl
HO
O
O
N
N
Bn

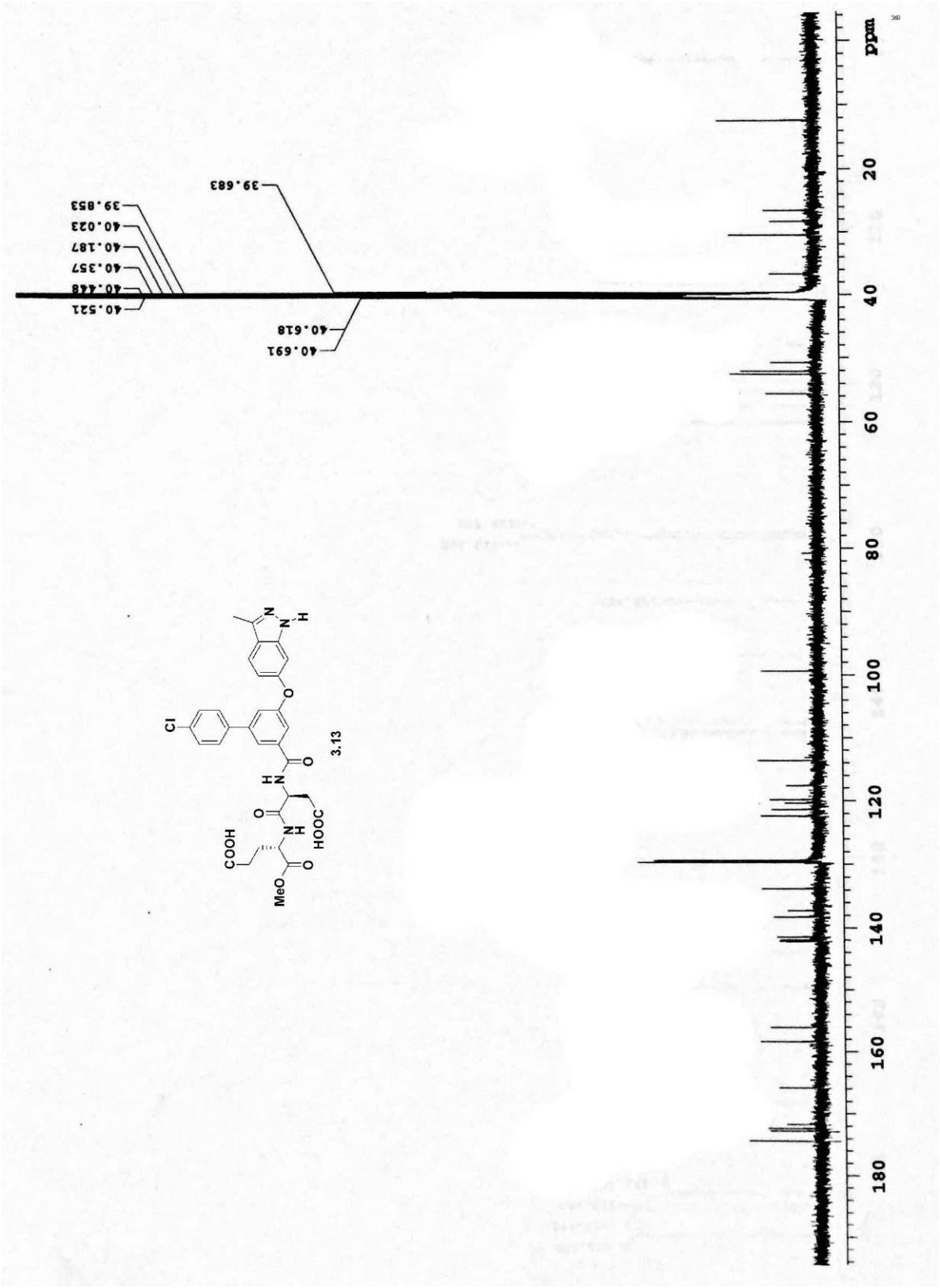
3.33b

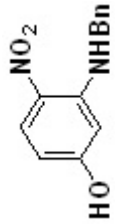

190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
f1 (ppm)


Cl
O
N
H
N
O
COOtBu
O
N
H
O
COOtBu
MeO
O
3.35b
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
f1 (ppm)
1.00
1.22
1.17
3.84
1.22
3.02
1.37
1.11
1.24
1.25
1.26
2.96
1.11
1.15
2.62
1.48
1.33
11.70
9.03

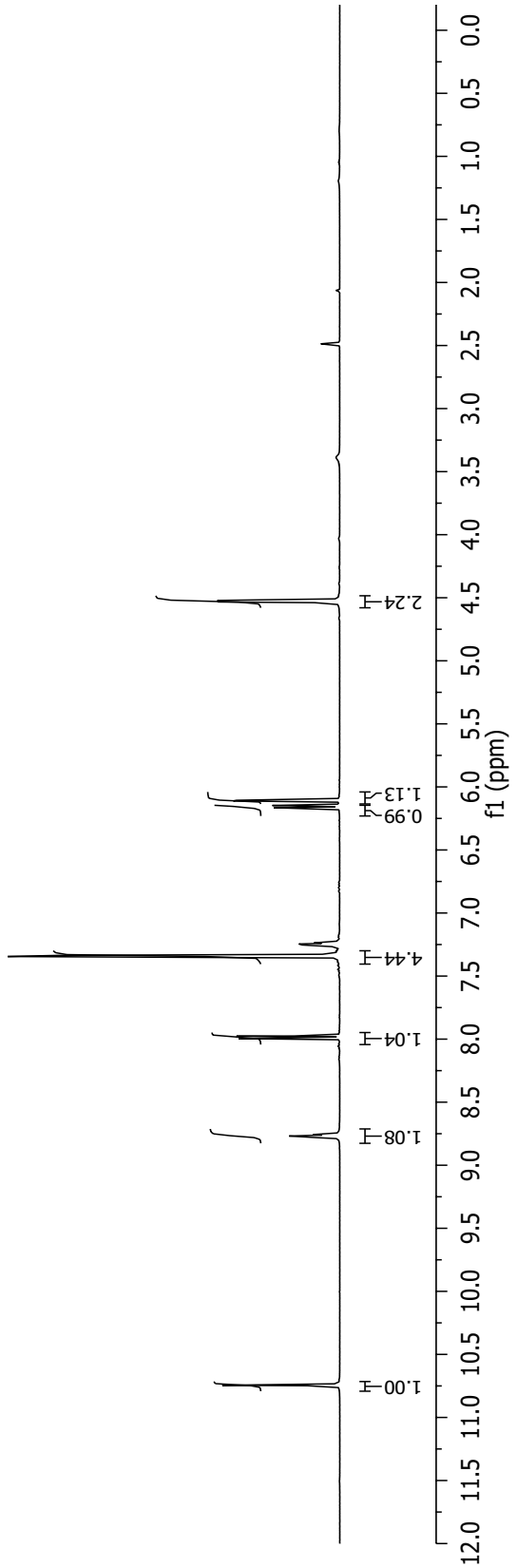

Cl
O
N
H
N
O
COOtBu
3.35b
N
H
O
N
H
COOtBu
MeO
O
0
10
20
30
40
50
60
70
80
90
100
110
120
130
140
150
160
170
180
190
f1 (ppm)

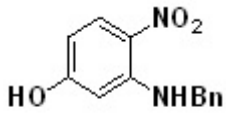

1.787
1.797
1.814
1.929
1.938
1.955
1.969
2.265
2.280
2.294
2.457
2.482
2.614
2.633
2.648
2.666
2.727
2.736
2.760
2.769
3.149
3.564
3.592
4.244
4.252
4.271
4.814
4.822
4.839
5.738
6.865
6.879
6.882
6.992
7.466
7.520
7.537
7.711
7.728
7.747
7.764
7.965
8.401
8.416
8.833
8.849
3.13
Cl
O
N
H
O
N
H
O
O
HOOC
COOH
MeO
ppm
1.01
3.40
0.99
0.94
0.98
0.96
3.29
0.84
0.90
1.17
1.10
3.00
1.07
2.16
1.14
1.08
0.55
1.36
1.86


3.13
39.683
40.691
40.618
40.521
40.428
40.357
40.187
40.023
39.853
ppm
20
40
60
80
100
120
140
160
180

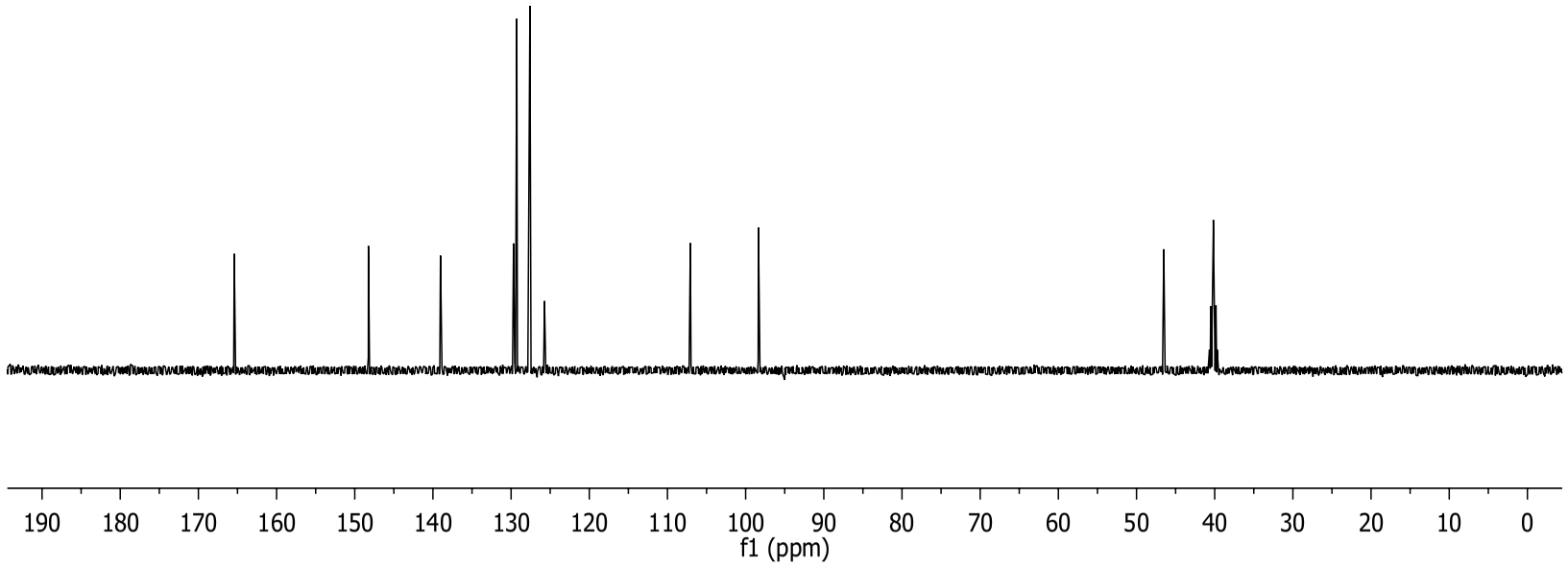


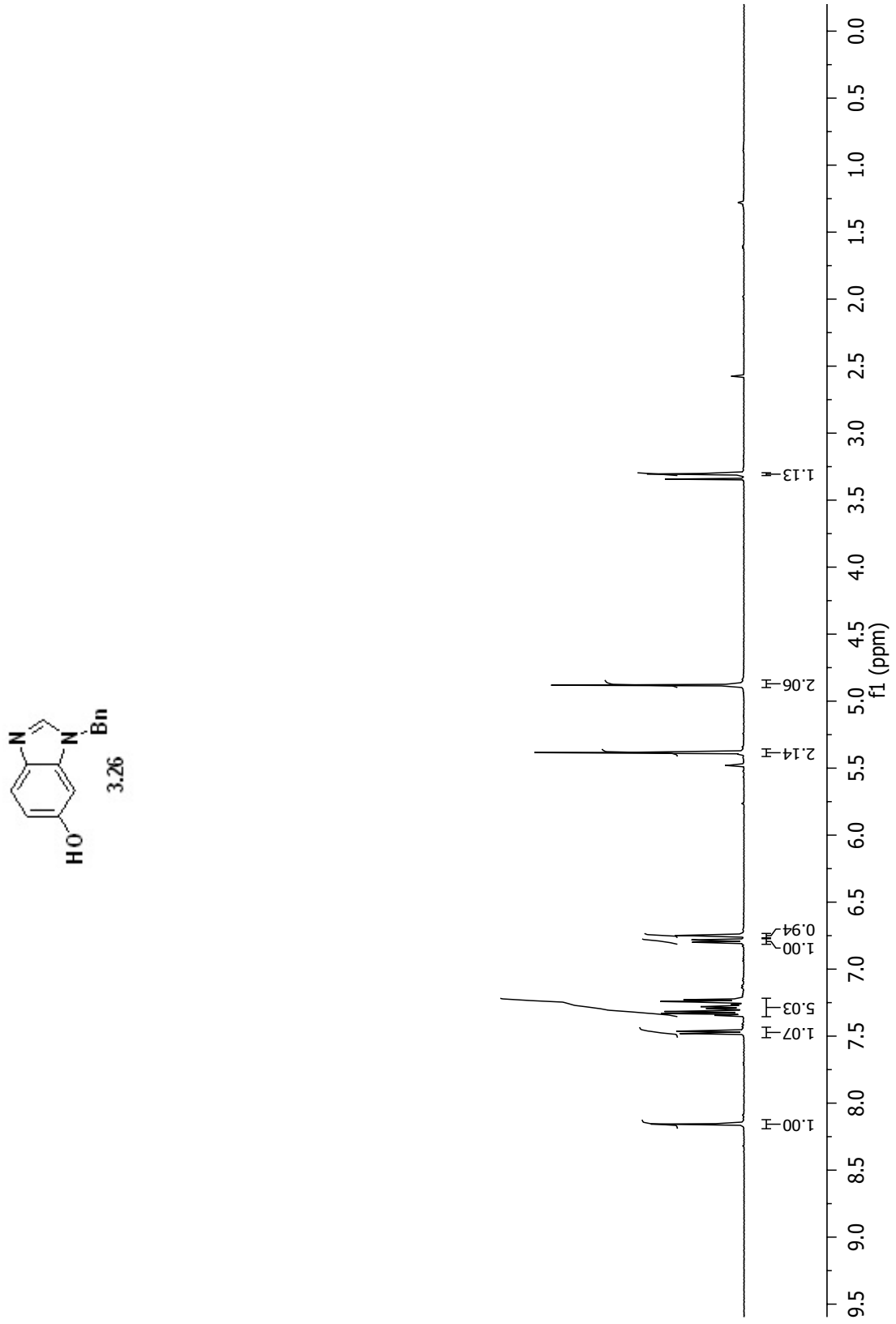
3.25

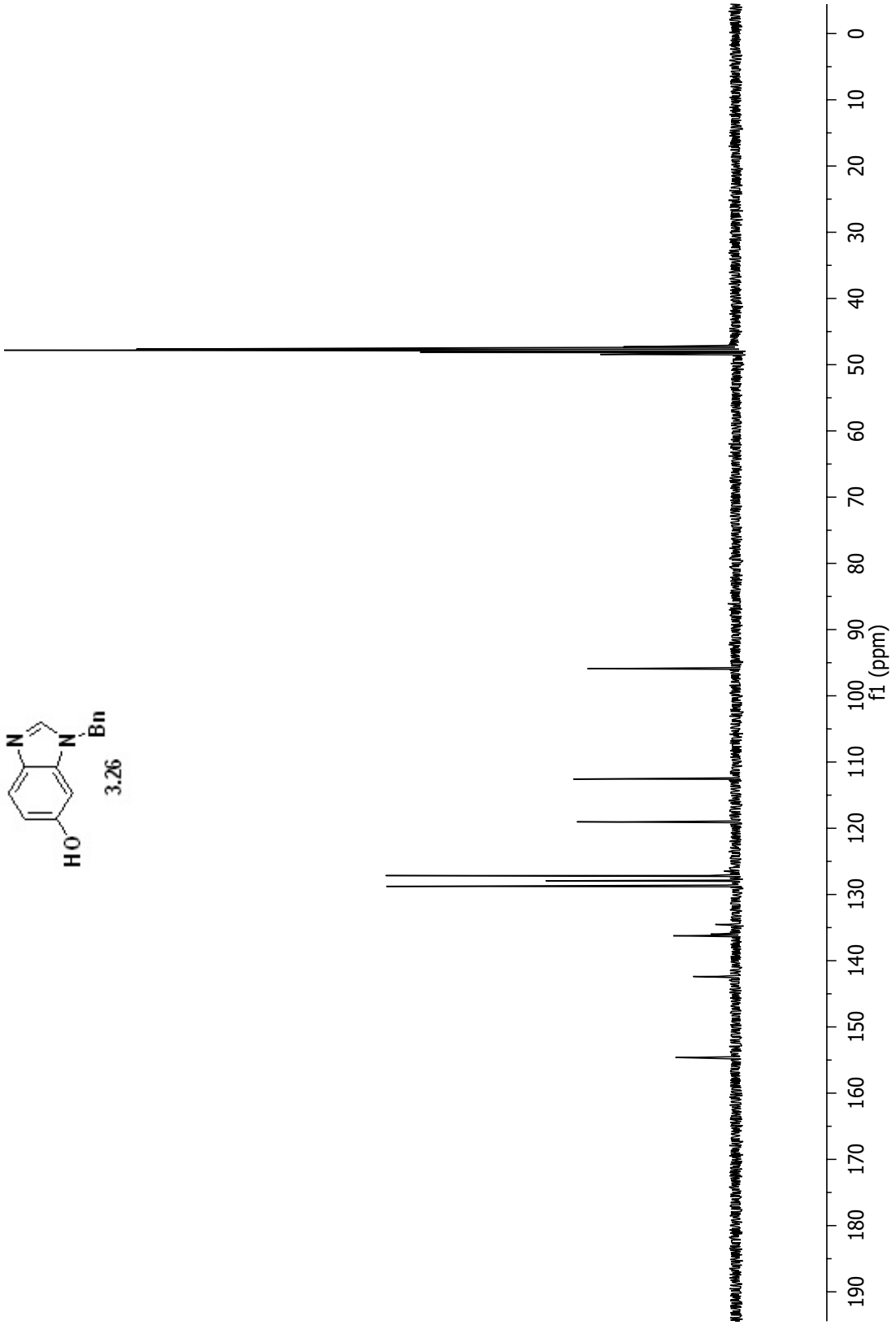


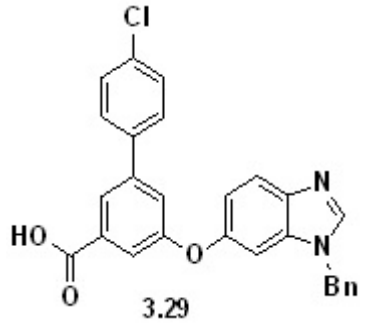

NO2
HO
NHBn

3.25

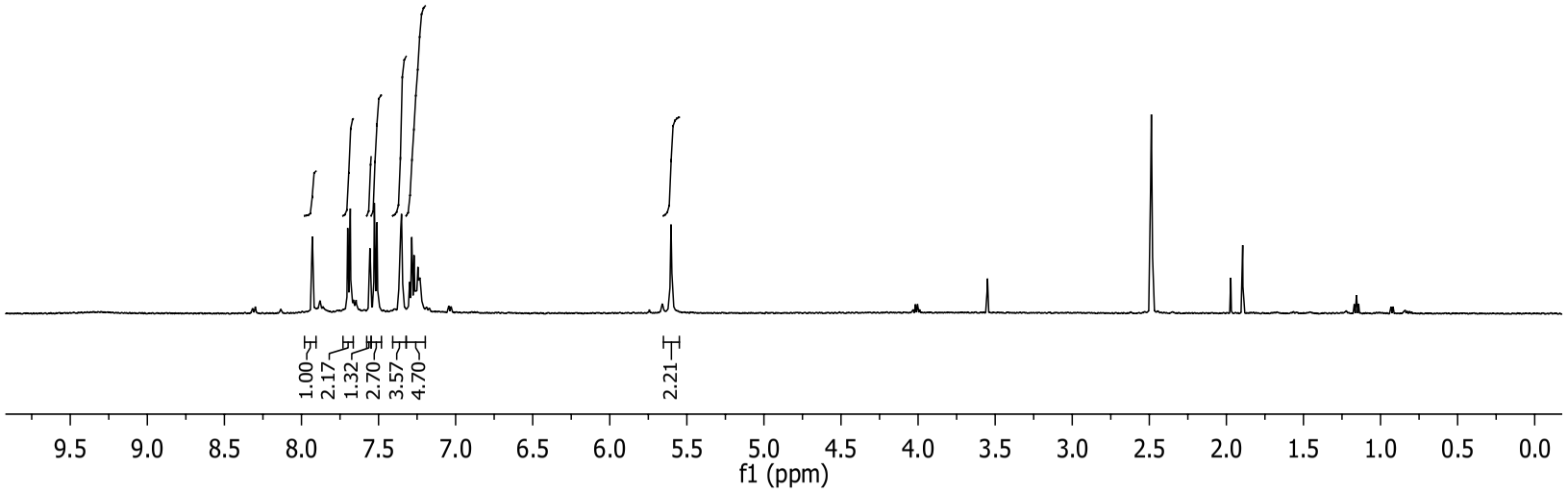

190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
f1 (ppm)

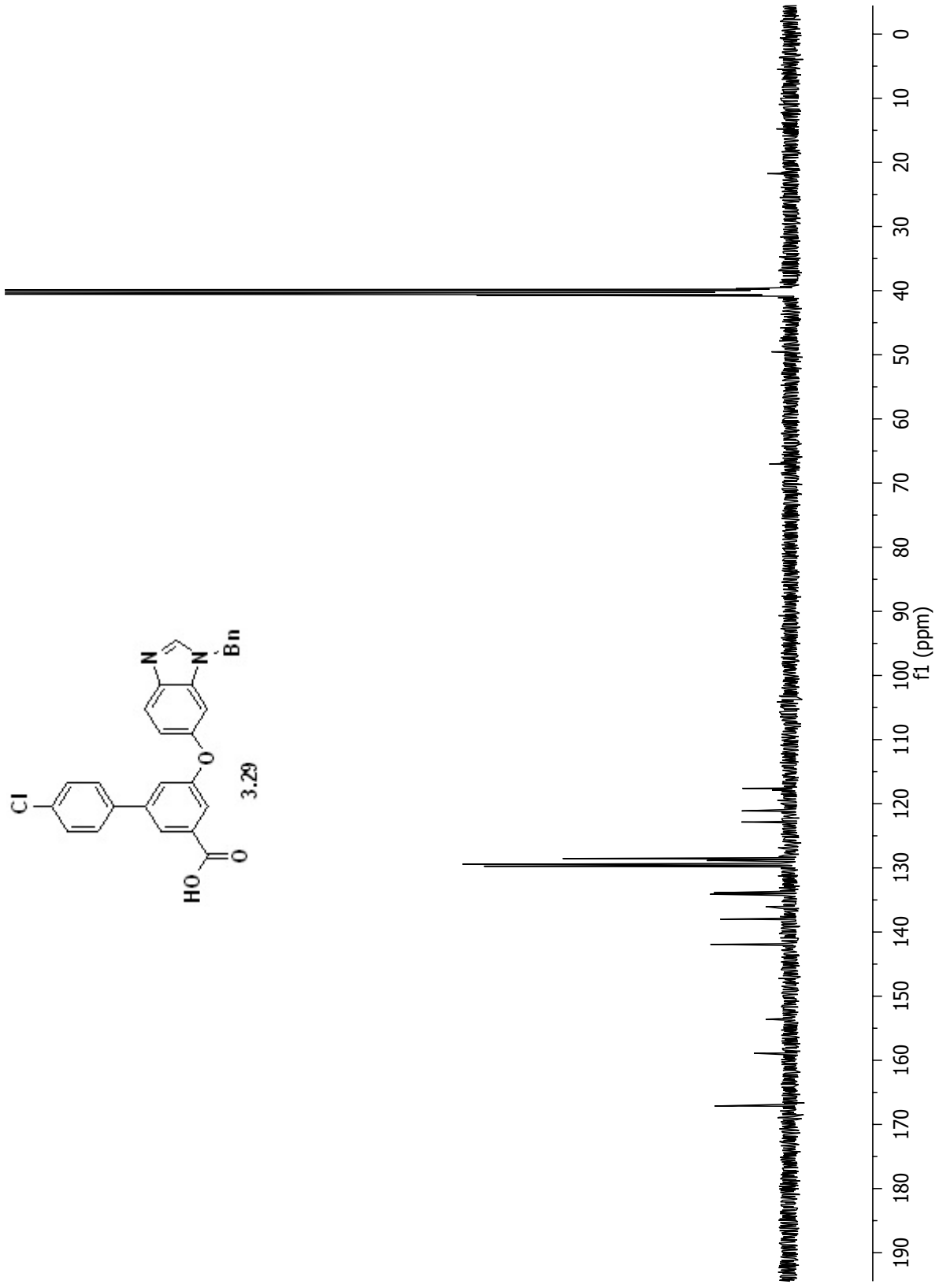

N
N
Bn
HO
3.26
1.13
2.06
2.14
1.00
0.94
5.03
1.07
1.00
9.5
9.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
f1 (ppm)

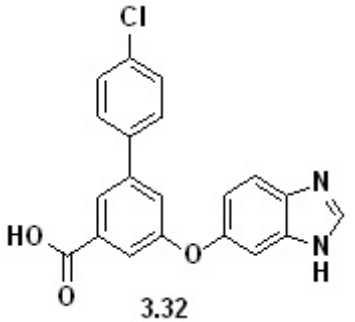

N
N
Bn
HO
3.26
0
10
20
30
40
50
60
70
80
90
100
110
120
130
140
150
160
170
180
190
f1 (ppm)


Cl
HO
O
O
N
N
Bn

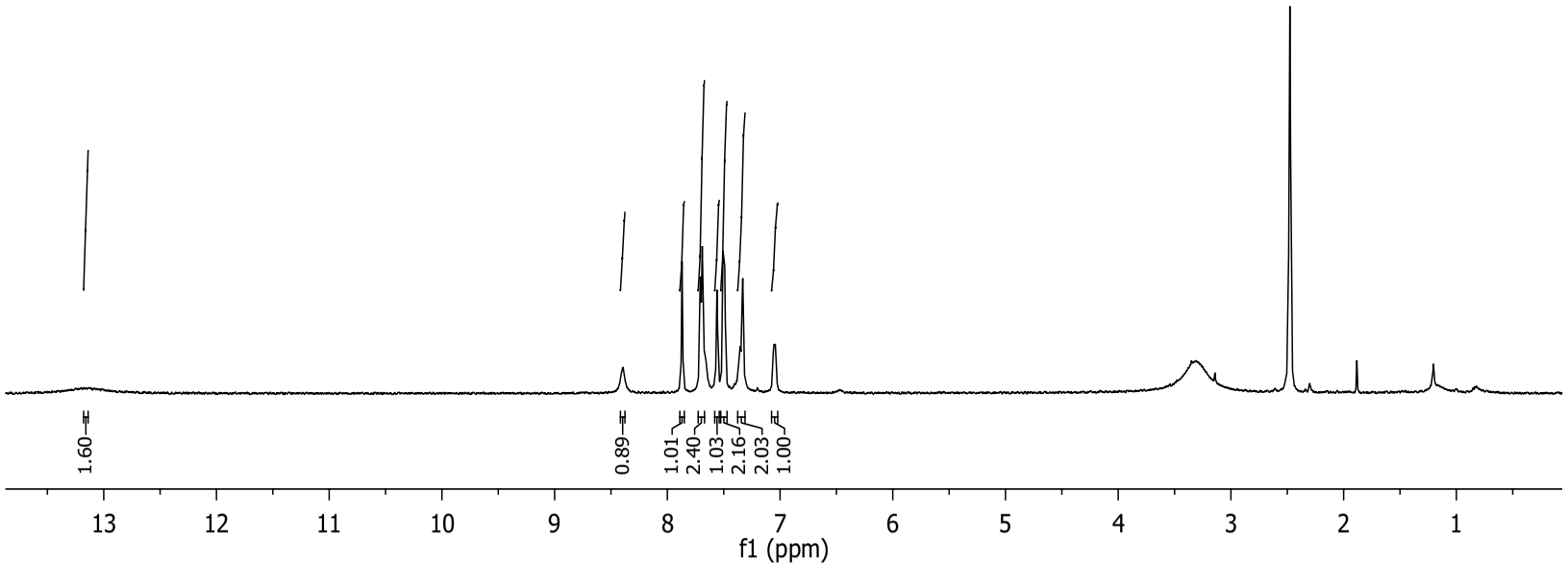
3.29

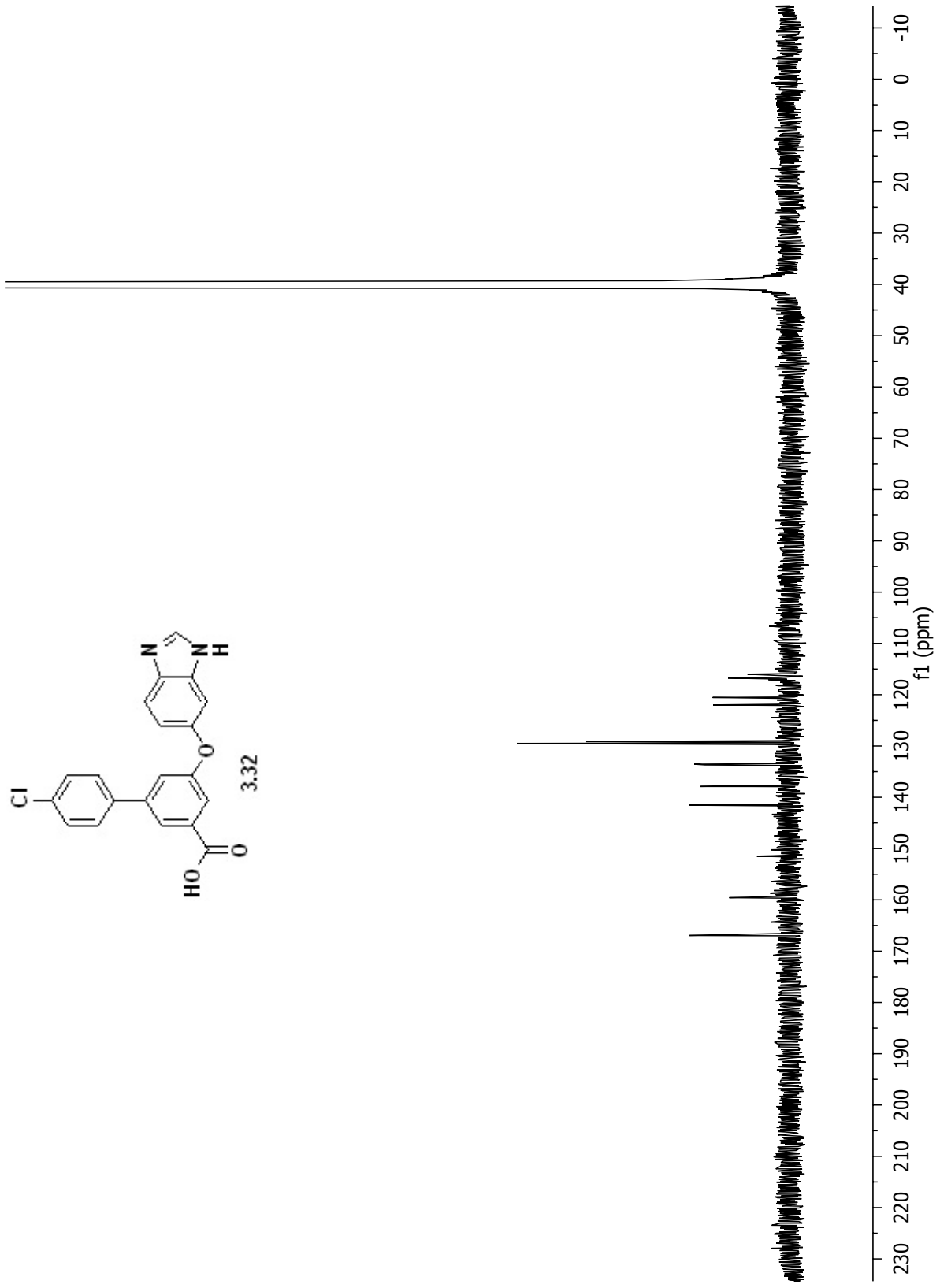

1.00
2.17
1.32
2.70
3.57
4.70
2.21
9.5 9.0 8.5 8.0 7.5 7.0 6.5 6.0 5.5 5.0 4.5 4.0 3.5 3.0 2.5 2.0 1.5 1.0 0.5 0.0
f1 (ppm)

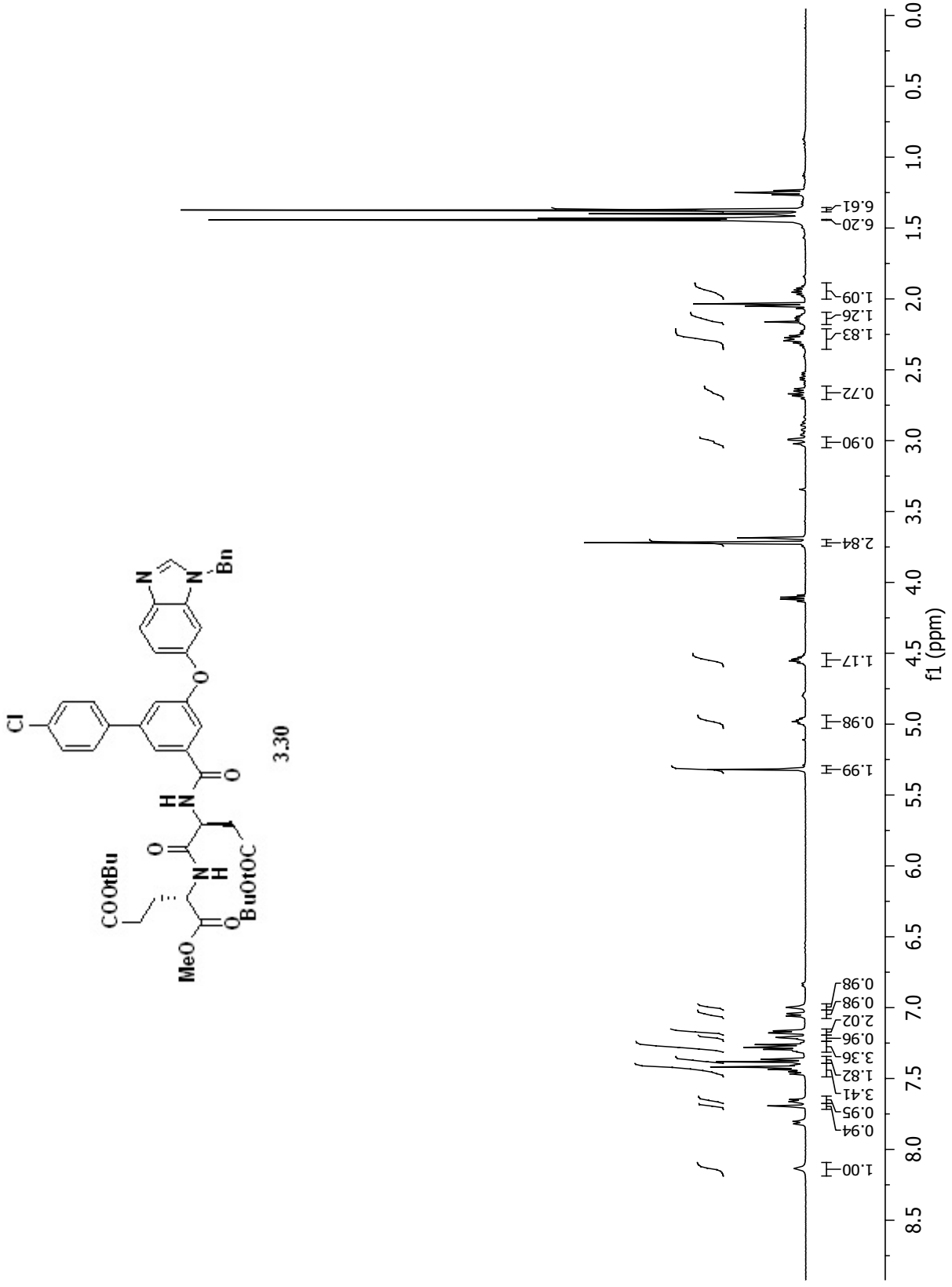

Cl
O
HO
O
N
N
Bn
3.29
f1 (ppm)
0
10
20
30
40
50
60
70
80
90
100
110
120
130
140
150
160
170
180
190

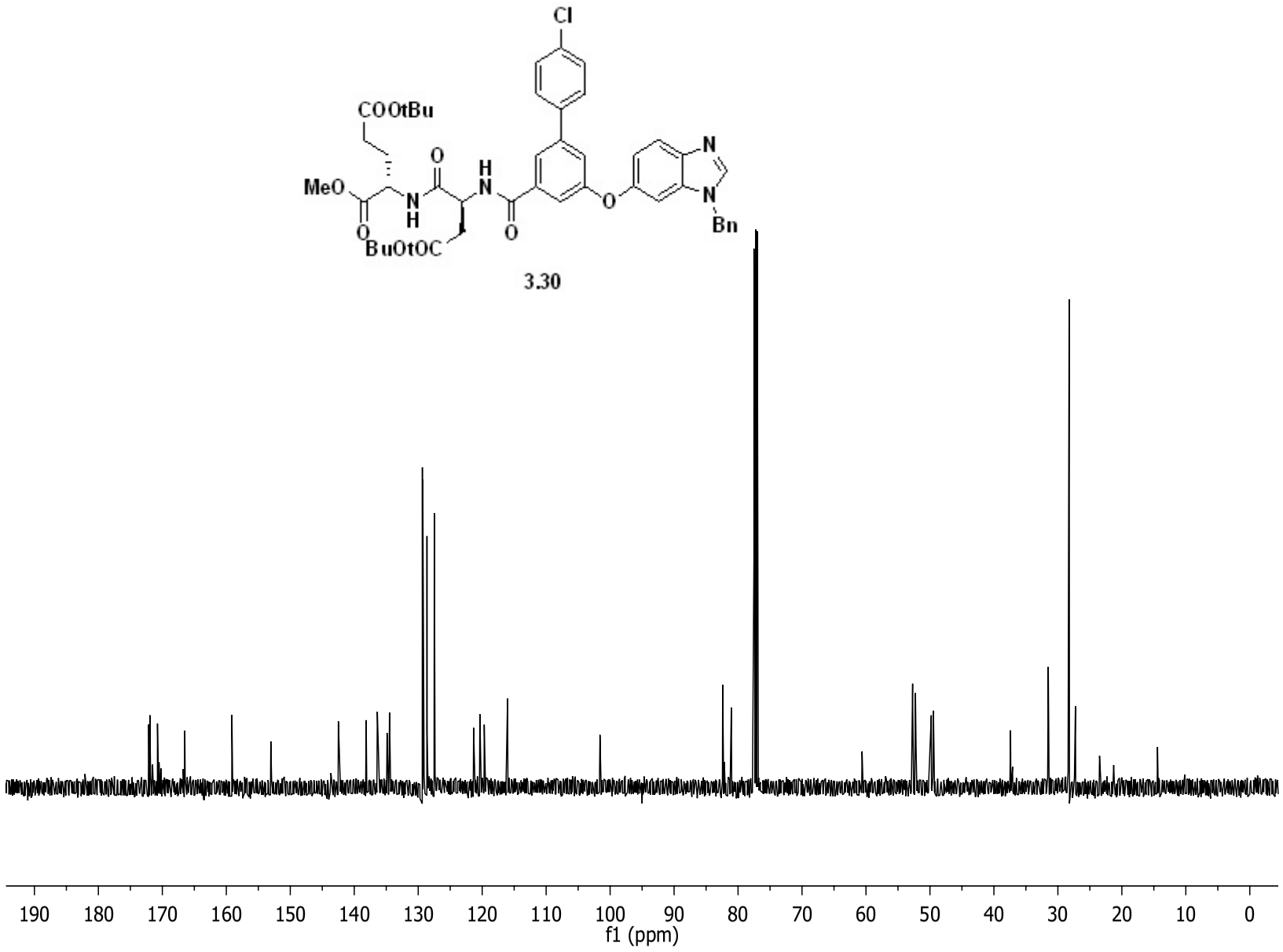

Cl
HO
O
O
N
N
H

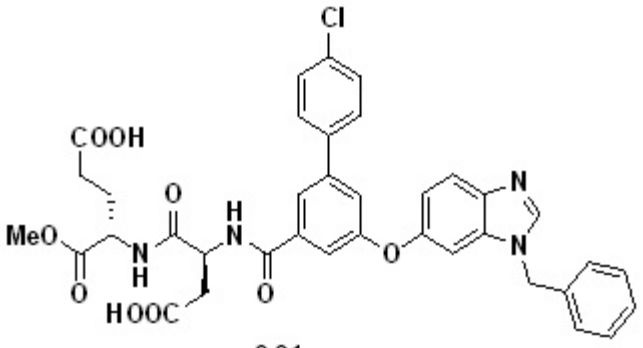
3.32


1.60
0.89
1.01
2.40
1.03
2.16
2.03
1.00
13
12
11
10
9
8
7
6
5
4
3
2
1
f1 (ppm)

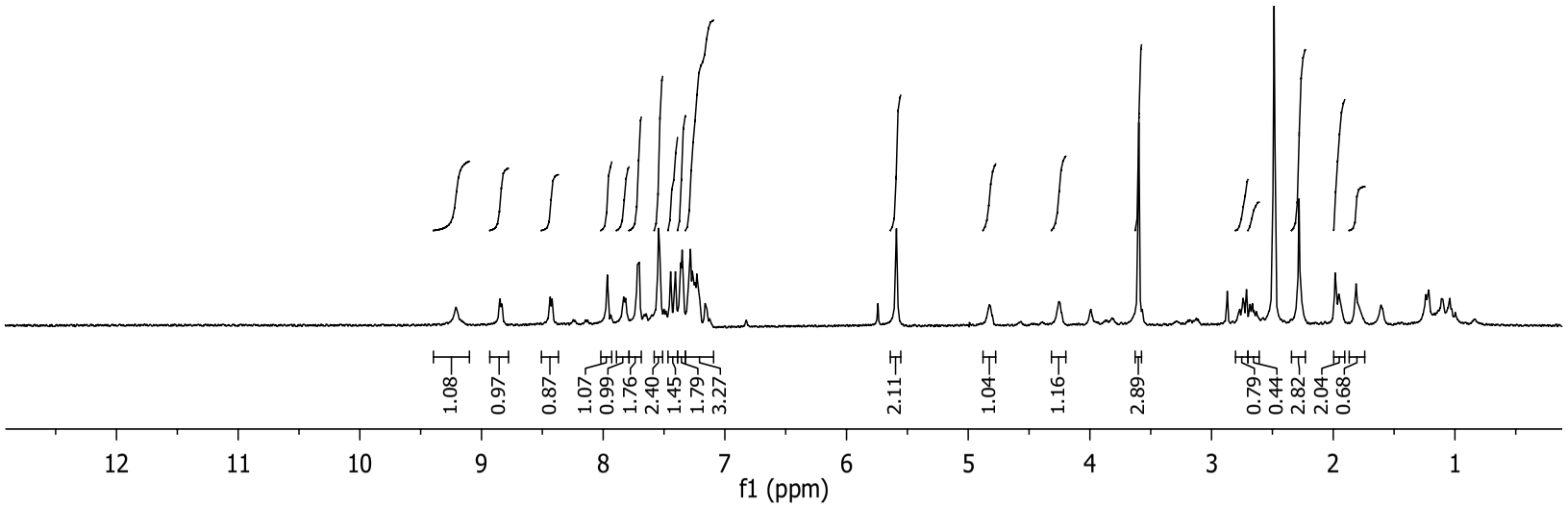

Cl
O
HO
O
N
N
H
3.32
-10
0
10
20
30
40
50
60
70
80
90
100
110
120
130
140
150
160
170
180
190
200
210
220
230
f1 (ppm)

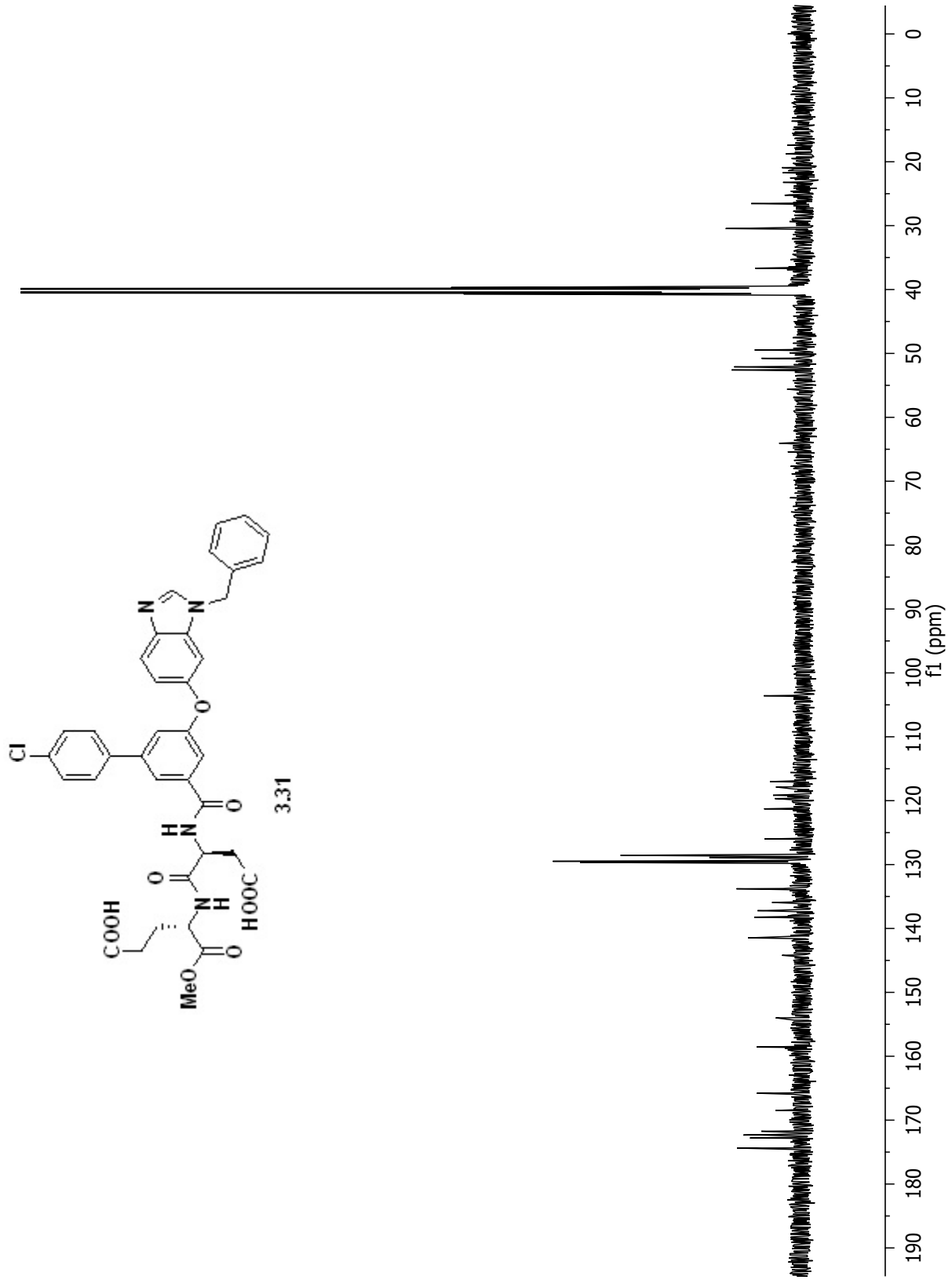

Cl
N
N
Bn
O
O
H
N
O
O
COOtBu
N
H
O
MeO
O
BuOtOC
3.30
f1 (ppm)
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
1.00
0.94
0.95
3.41
1.82
3.36
0.96
2.02
0.98
0.98
1.99
1.17
0.98
2.84
0.90
0.72
1.83
1.26
1.09
6.20
6.61

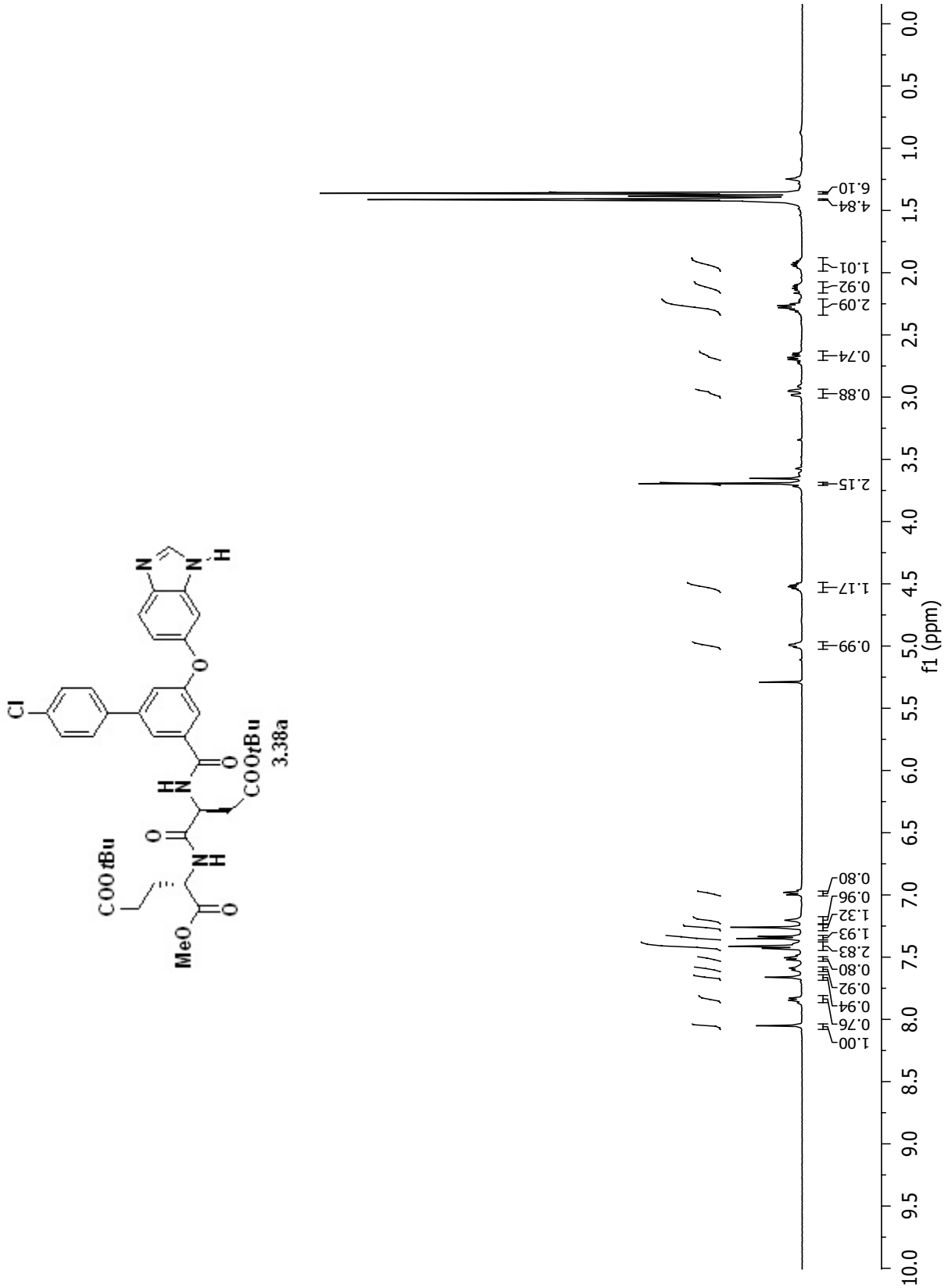

Cl
COOtBu
MeO
O
H
N
N
H
O
O
O
BuOtOC
O
N
N
Bn
3.30
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
f1 (ppm)

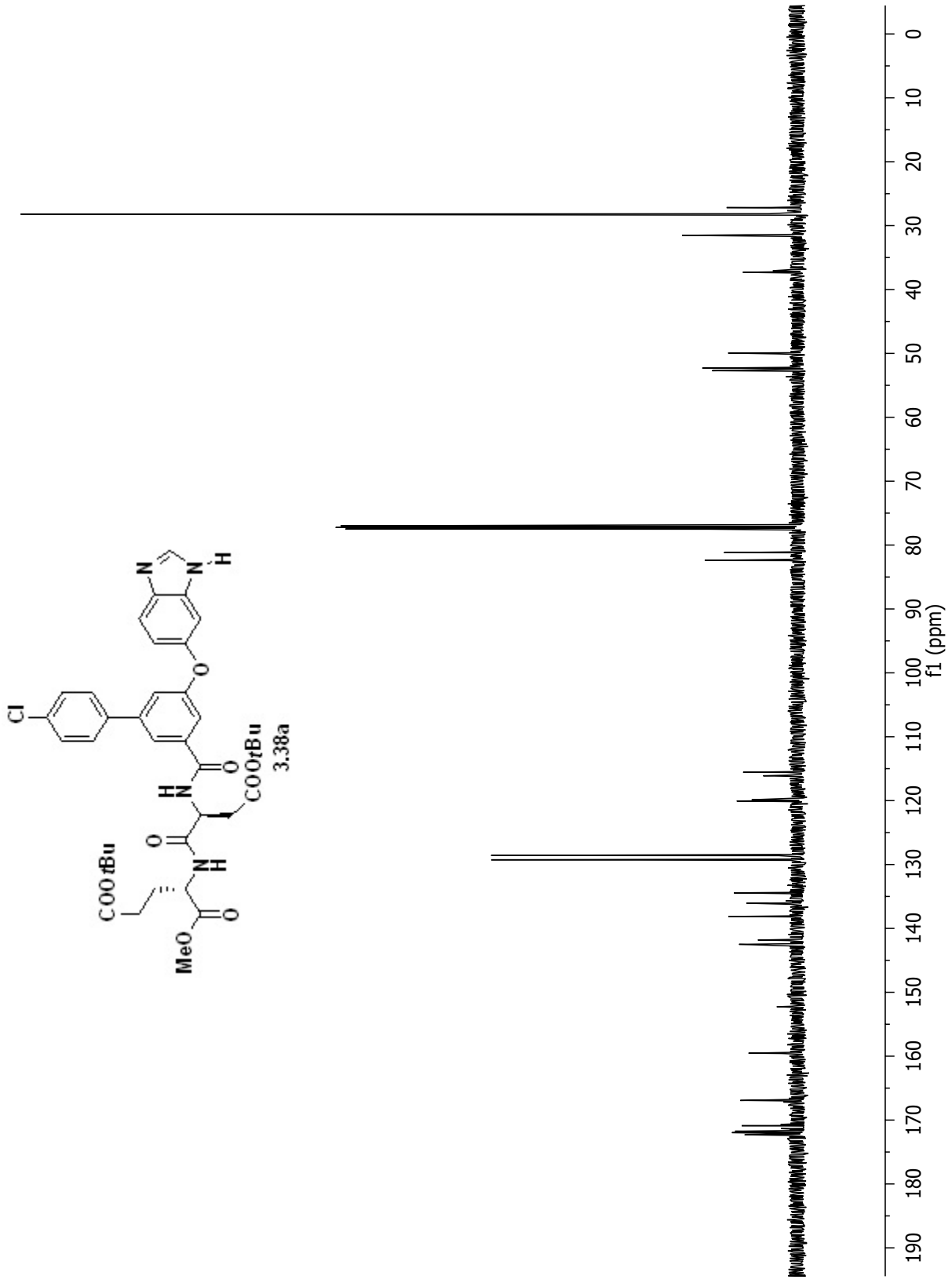

Cl
COOH
MeO
O
N
H
O
H
N
O
HOOC
O
N
N
O

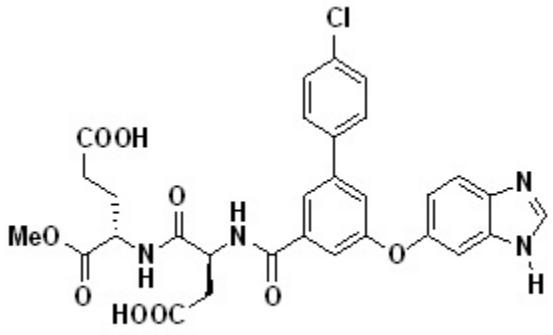
3.31


1.08
0.97
0.87
1.07
0.99
1.76
2.40
1.45
1.79
3.27
2.11
1.04
1.16
2.89
0.79
0.44
2.82
2.04
0.68
12
11
10
9
8
7
6
5
4
3
2
1
f1 (ppm)

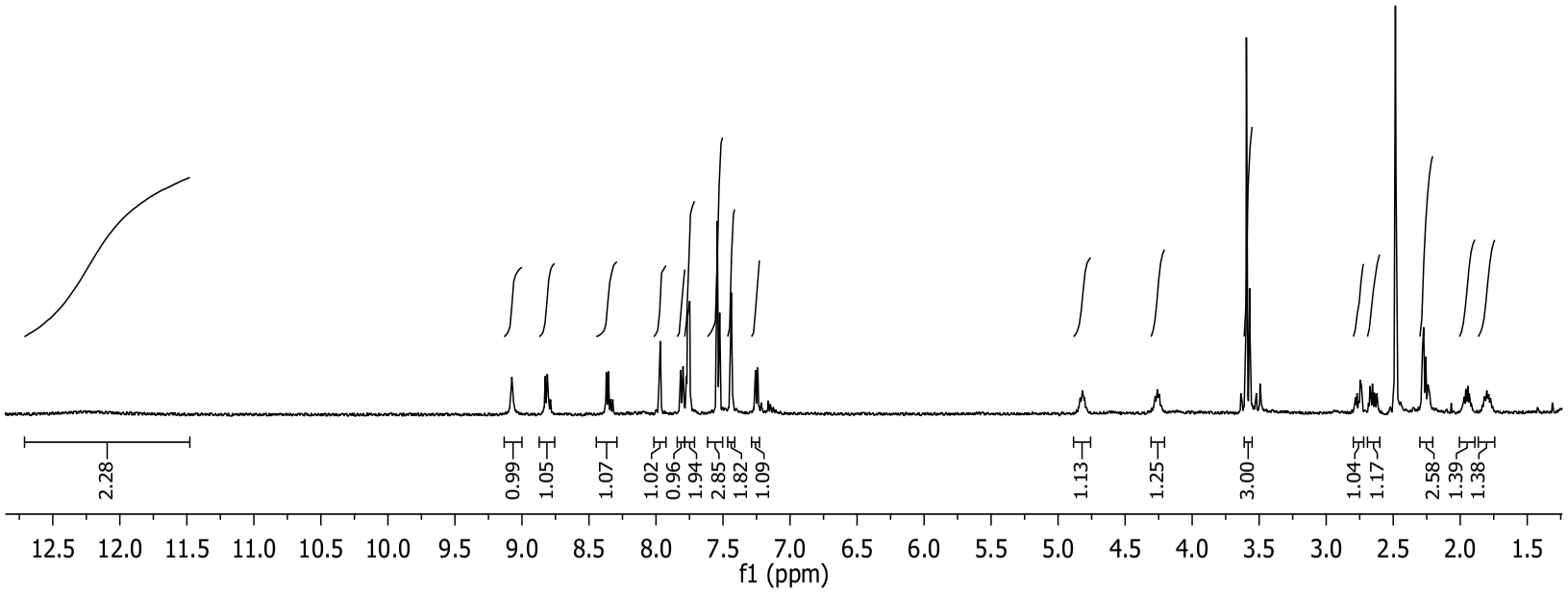

Cl
O
N
N
O
H
N
O
O
COOH
N
H
HOOC
O
MeO
3.31
f1 (ppm)
0
10
20
30
40
50
60
70
80
90
100
110
120
130
140
150
160
170
180
190

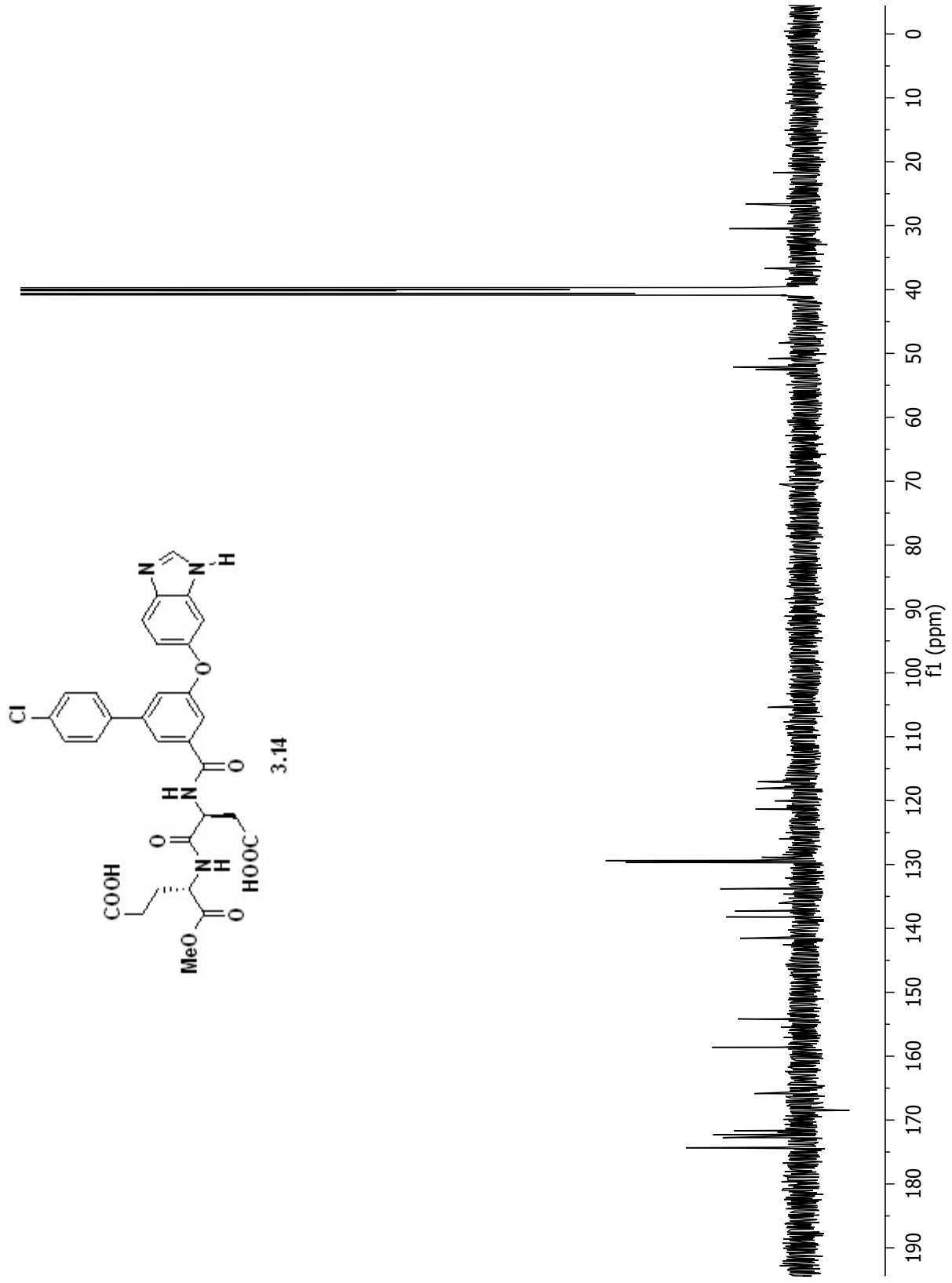

Cl
COO tBu
MeO
COO tBu
3.38a
f1 (ppm)
10.0
9.5
9.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
1.00
0.76
0.94
0.92
0.80
2.83
1.93
1.32
0.96
0.80
0.99
1.17
2.15
0.88
0.74
2.09
0.92
1.01
4.84
6.10

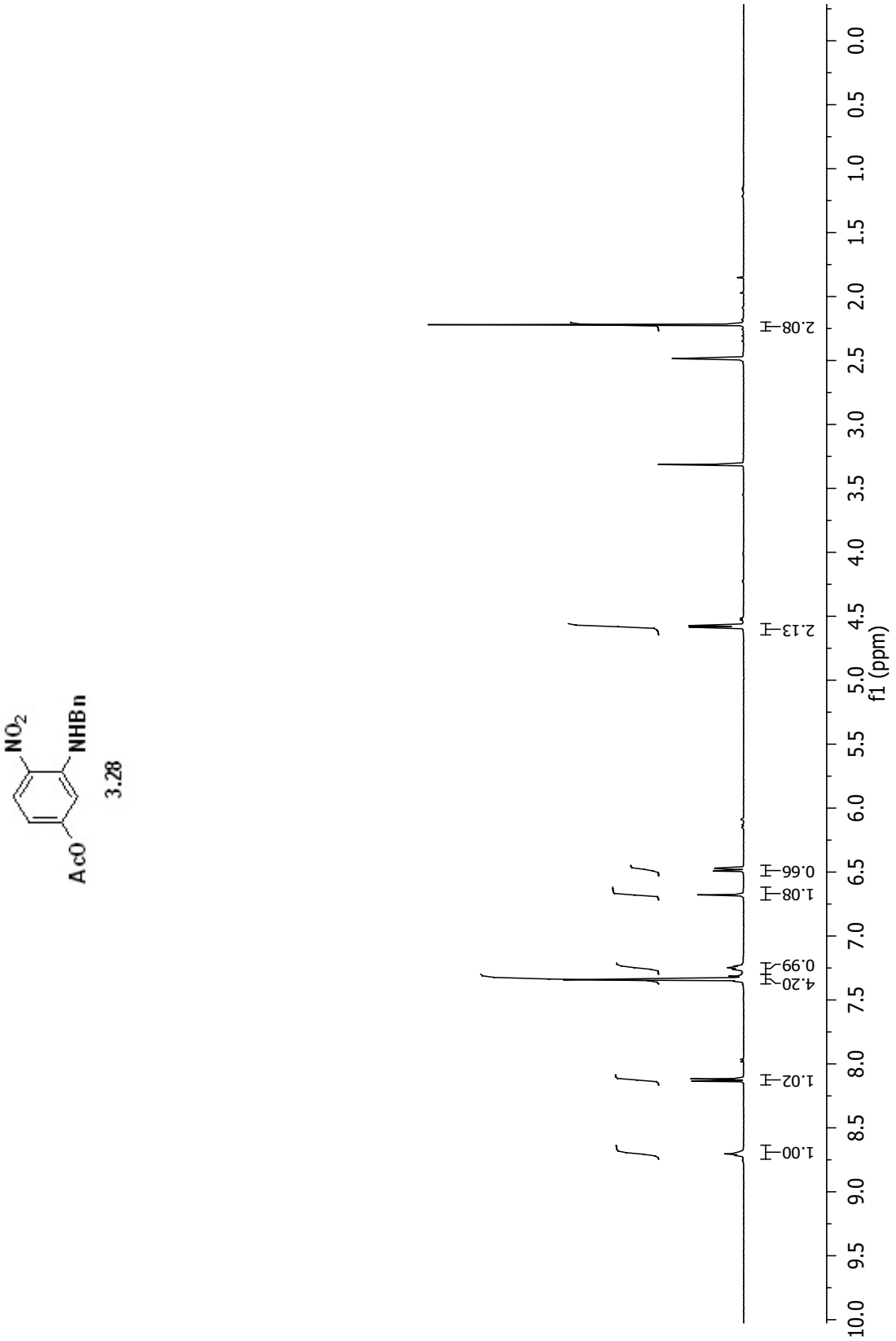

Cl
O
N
H
N
H
O
COOtBu
3.38a
N
H
O
COOtBu
O
MeO
0
10
20
30
40
50
60
70
80
90
100
110
120
130
140
150
160
170
180
190
f1 (ppm)

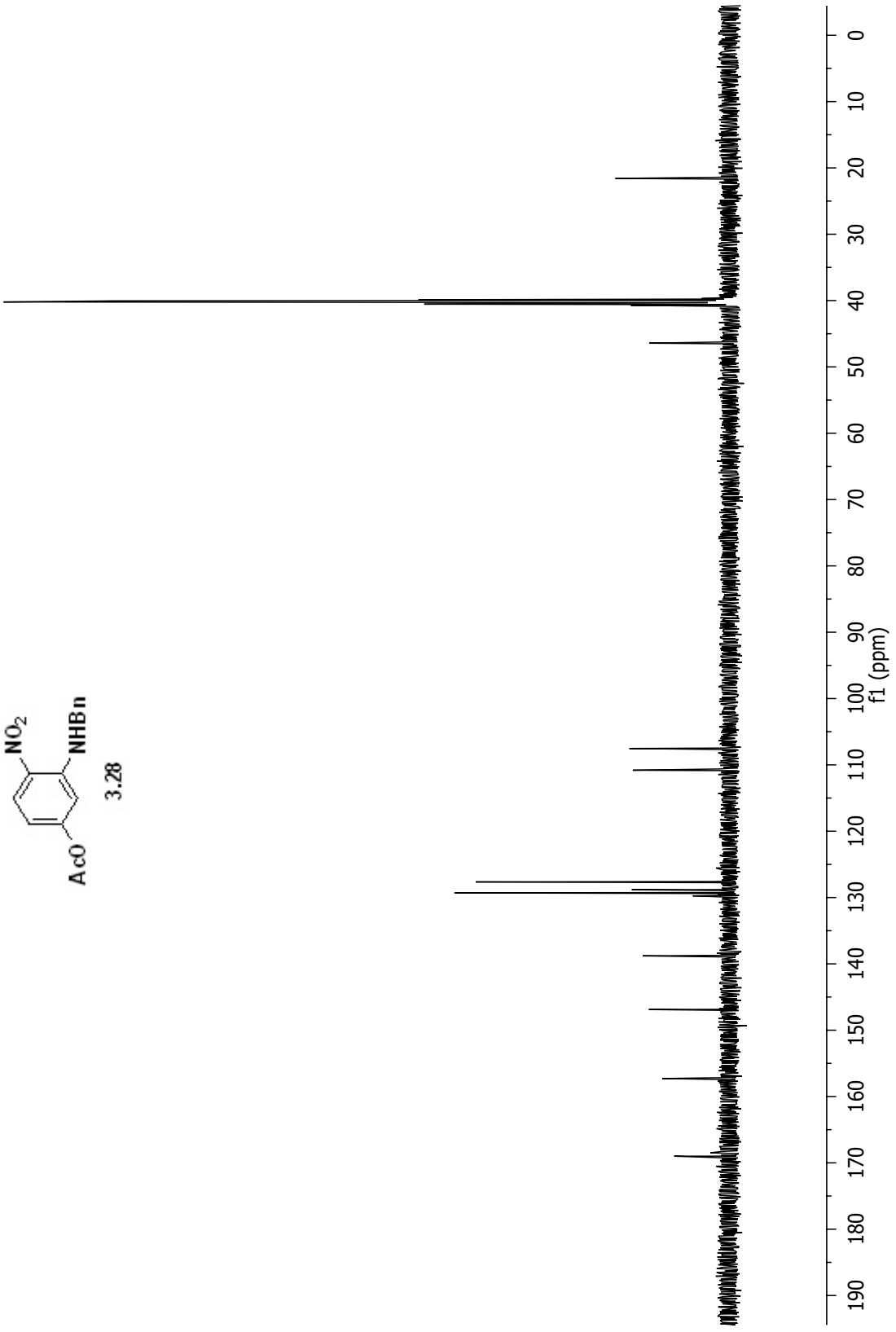

Cl
COOH
MeO
O
N
H
O
H
N
O
HOOC
O
N
N
H
O

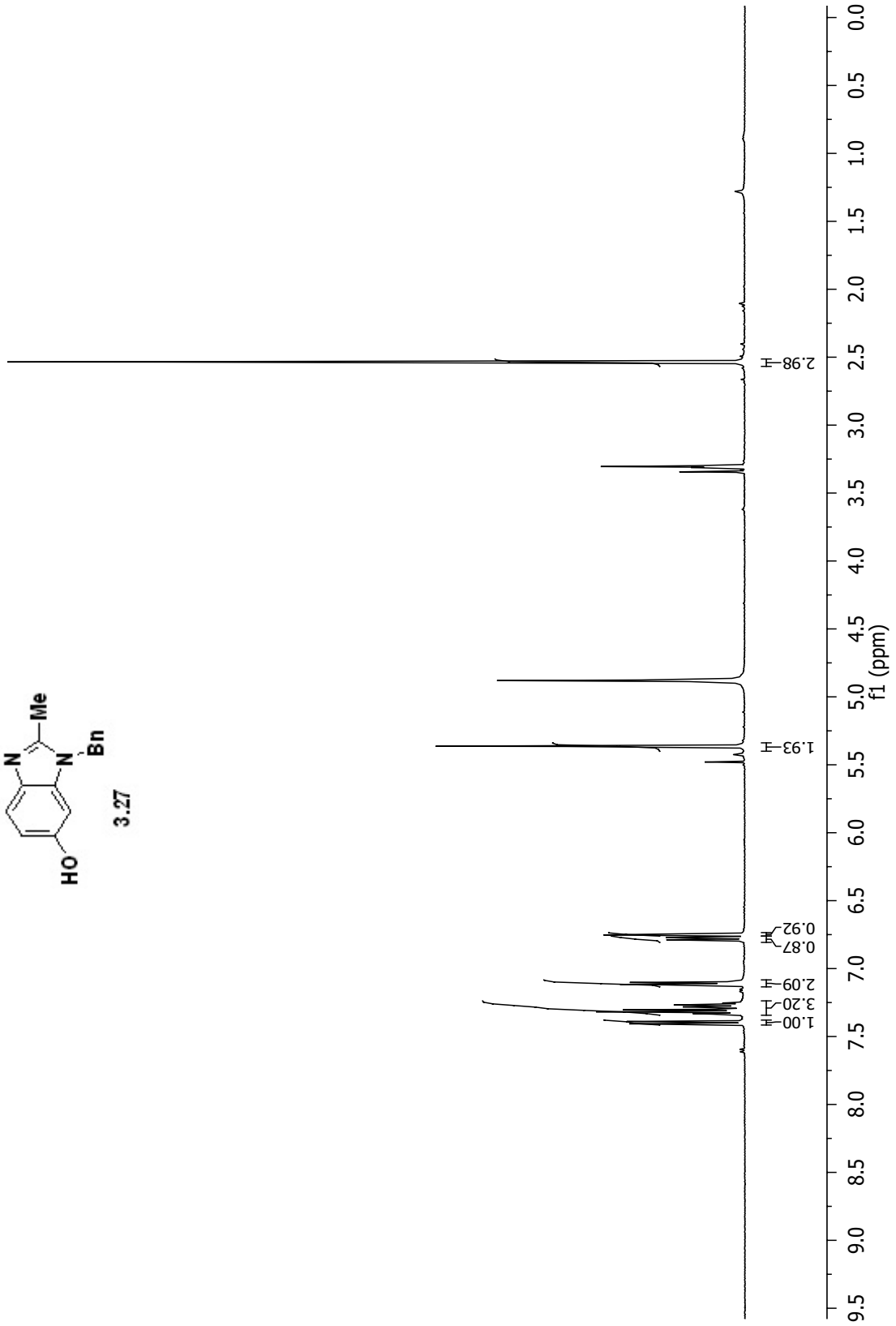
3.14

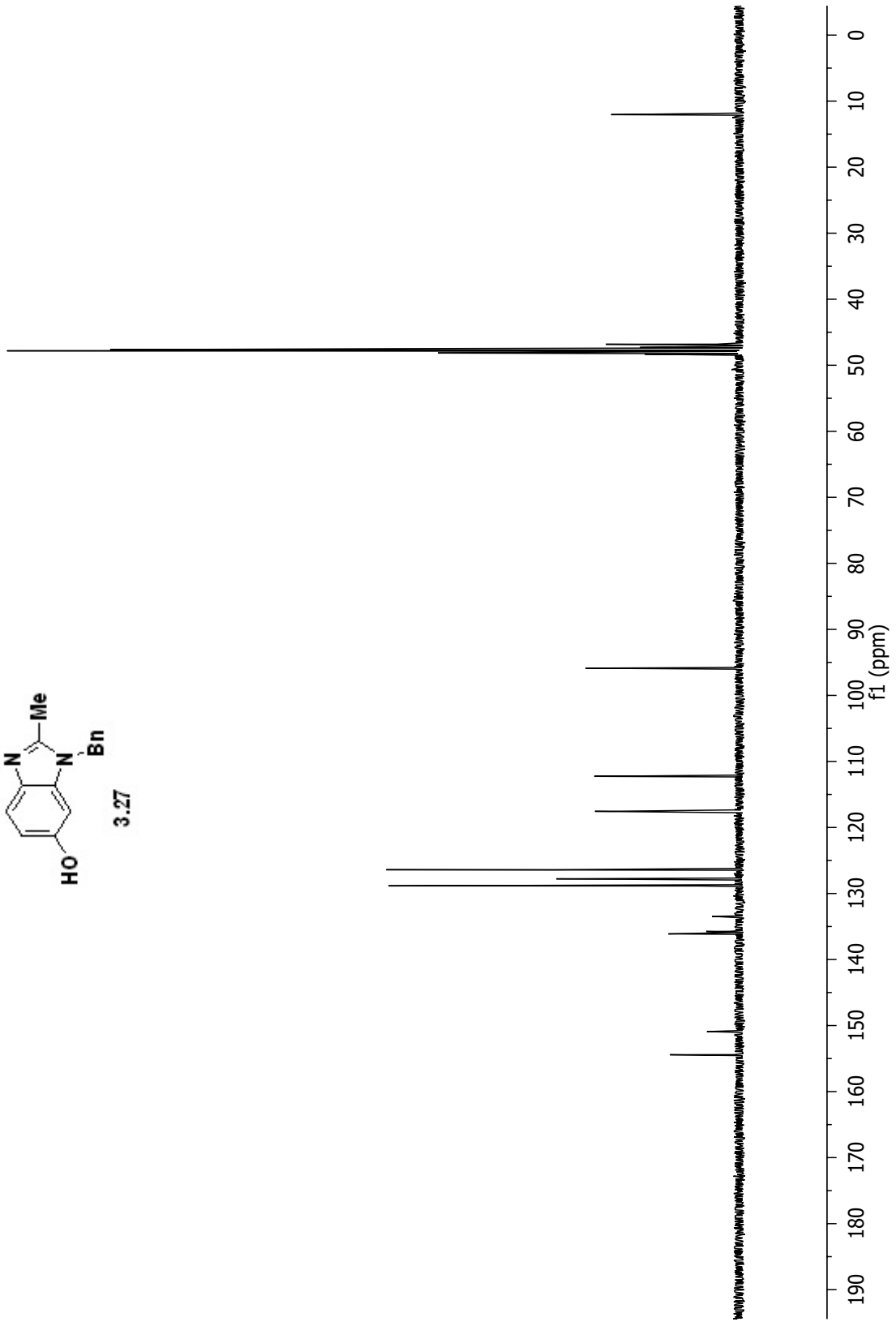

2.28
0.99
1.05
1.07
1.02
0.96
1.94
2.85
1.82
1.09
1.13
1.25
3.00
1.04
1.17
2.58
1.39
1.38
12.5 12.0 11.5 11.0 10.5 10.0 9.5 9.0 8.5 8.0 7.5 7.0 6.5 6.0 5.5 5.0 4.5 4.0 3.5 3.0 2.5 2.0 1.5
f1 (ppm)

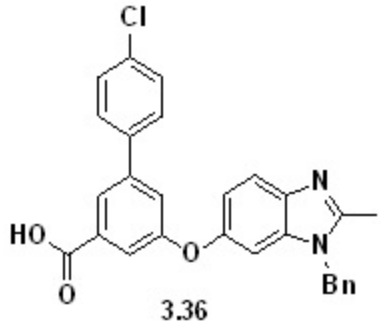

Cl
O
N
H
O
N
H
O
COOH
MeO
O
N
H
HOOC
O
3.14
f1 (ppm)
0
10
20
30
40
50
60
70
80
90
100
110
120
130
140
150
160
170
180
190


NO2
NHBn
AcO
3.28
2.08
2.13
0.66
1.08
4.20
0.99
1.02
1.00
10.0
9.5
9.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
f1 (ppm)

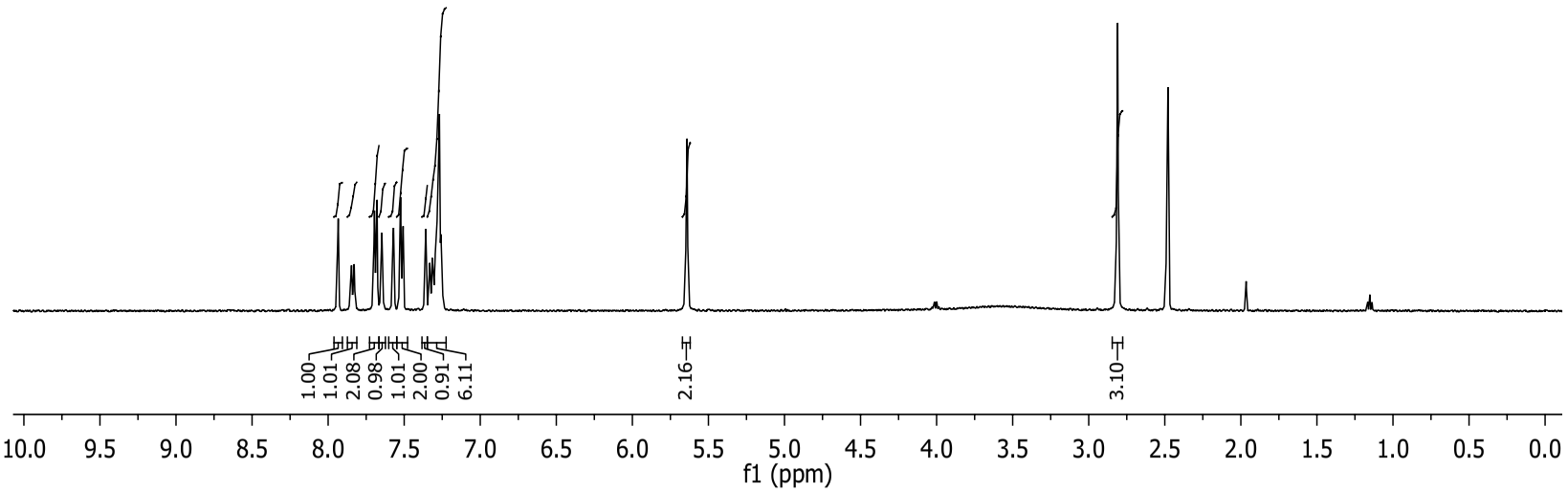

NO2
NHBn
AcO
3.28
f1 (ppm)

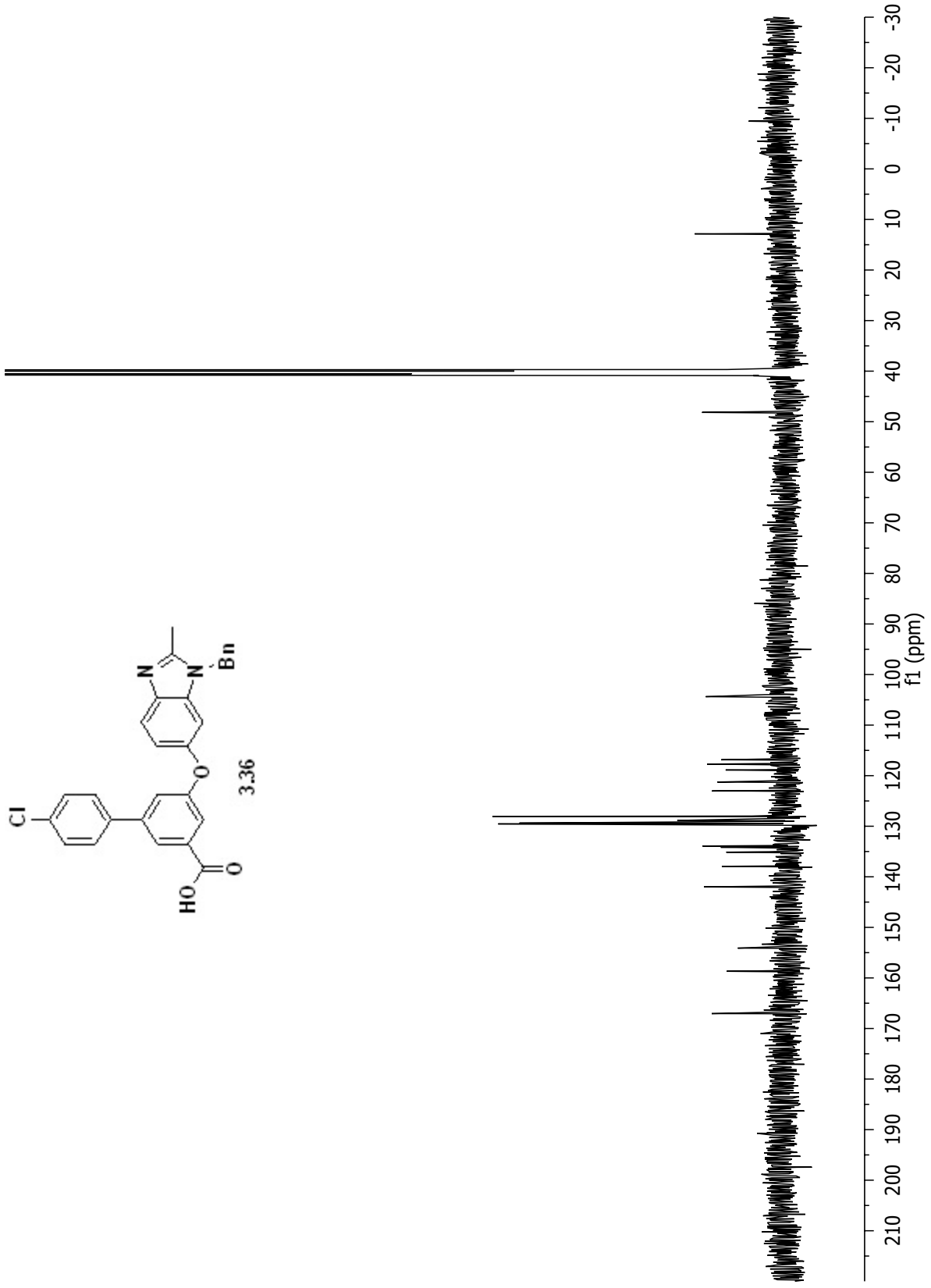

Me
N
N
Bn
HO
3.27
2.98
1.93
0.87
0.92
2.09
3.20
1.00
f1 (ppm)

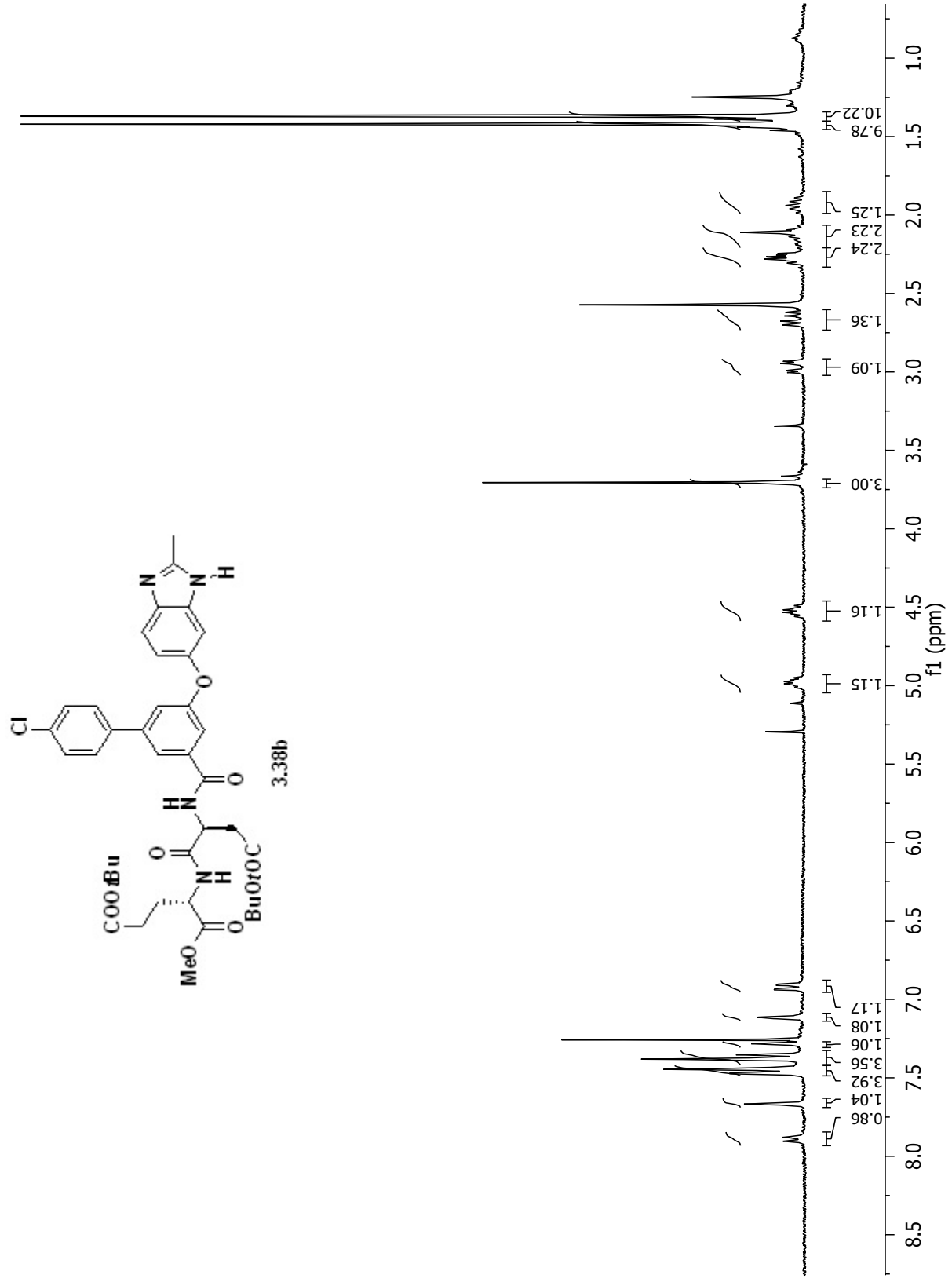

Me
N
N
Bn
HO
3.27
0
10
20
30
40
50
60
70
80
90
100
110
120
130
140
150
160
170
180
190
f1 (ppm)

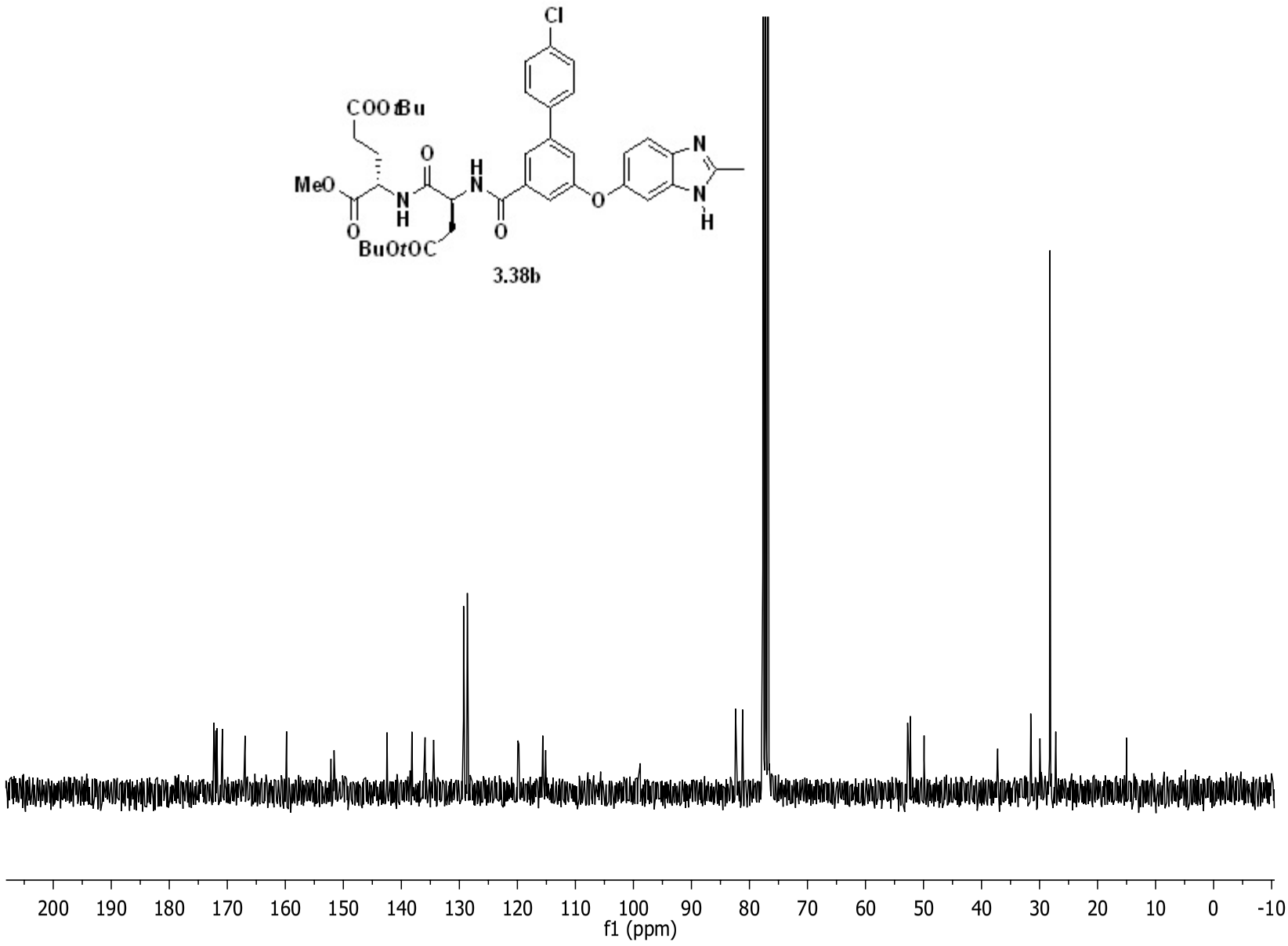

Cl
HO
O
O
N
N
Bn

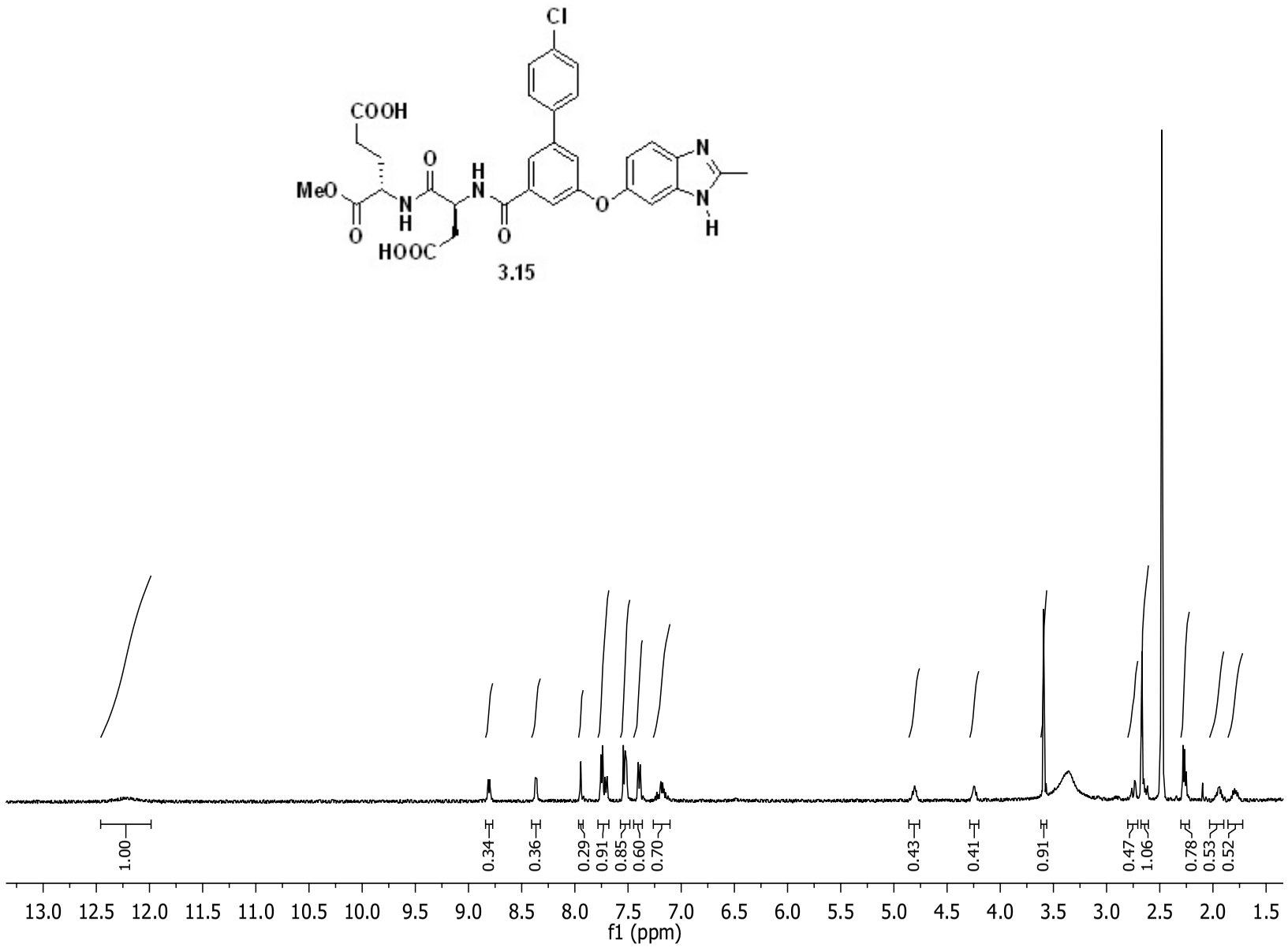
3.36

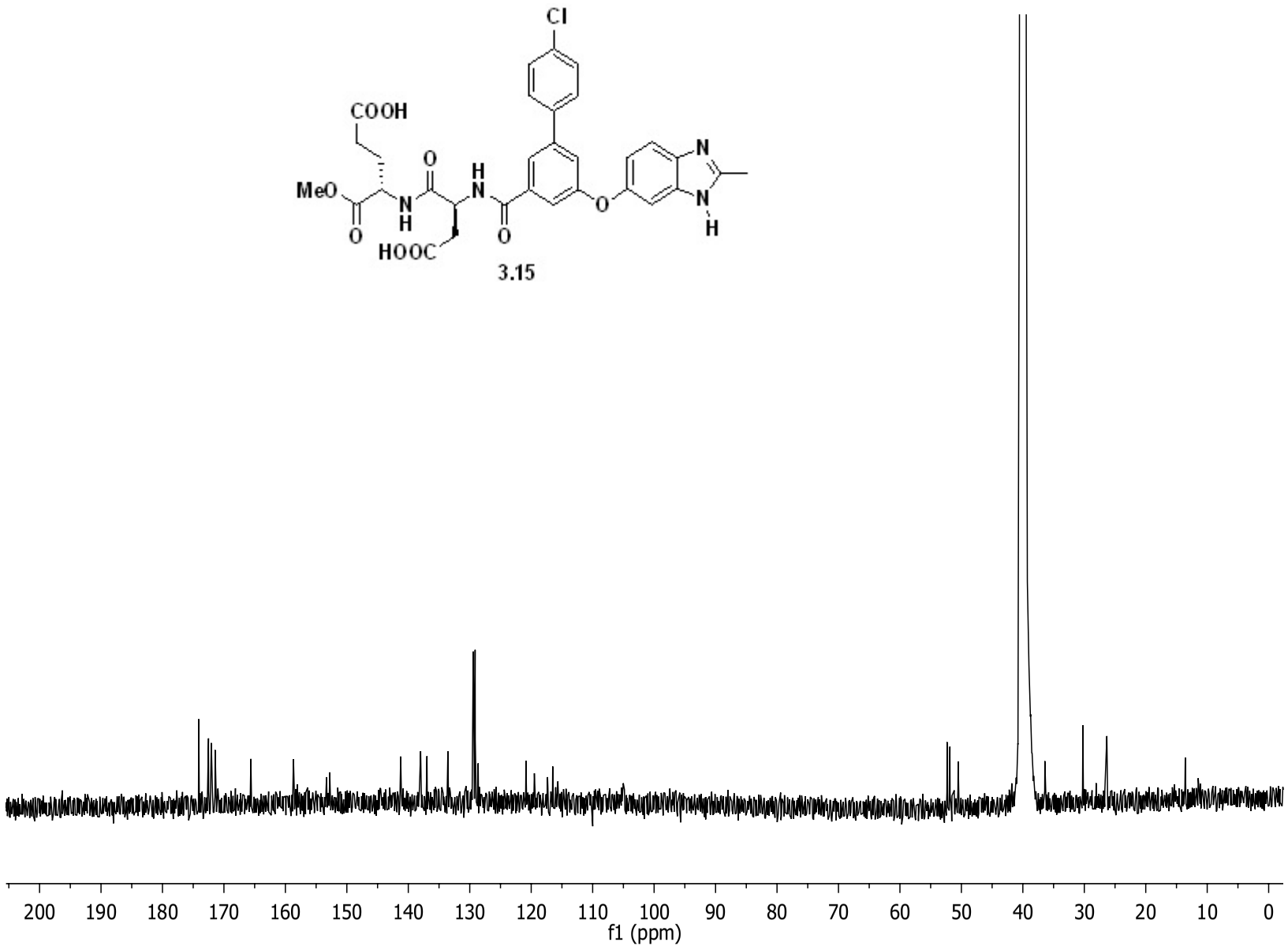

1.00
1.01
2.08
0.98
1.01
2.00
0.91
6.11
2.16
3.10
10.0 9.5 9.0 8.5 8.0 7.5 7.0 6.5 6.0 5.5 5.0 4.5 4.0 3.5 3.0 2.5 2.0 1.5 1.0 0.5 0.0
f1 (ppm)

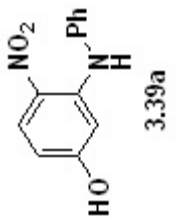

Cl
O
HO
O
N
N
Bn
3.36
f1 (ppm)


Cl
COOtBu
MeO
BuOtOC
3.38b
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
f1 (ppm)
0.86
1.04
3.92
3.56
1.06
1.08
1.17
1.15
1.16
3.00
1.09
1.36
2.24
2.23
1.25
9.78
10.22

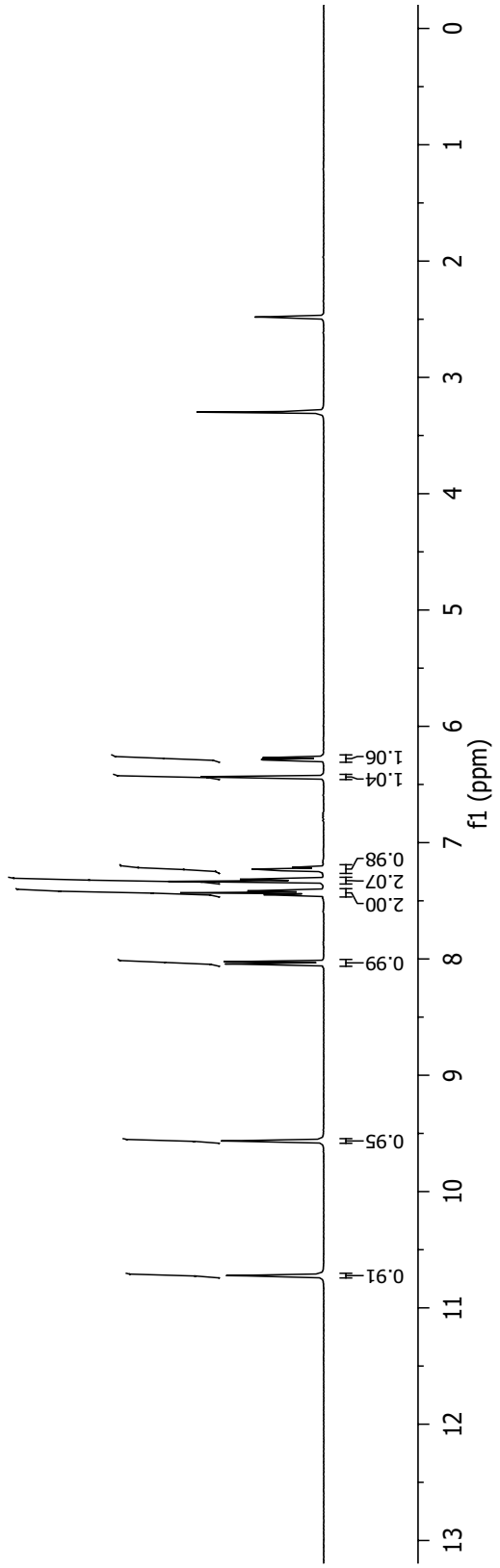

Cl
COOtBu
MeO
O
N
H
H
N
O
O
N
N
H
O
BuOtOC
3.38b
200 190 180 170 160 150 140 130 120 110 100 90 80 70 60 50 40 30 20 10 0 -10
f1 (ppm)

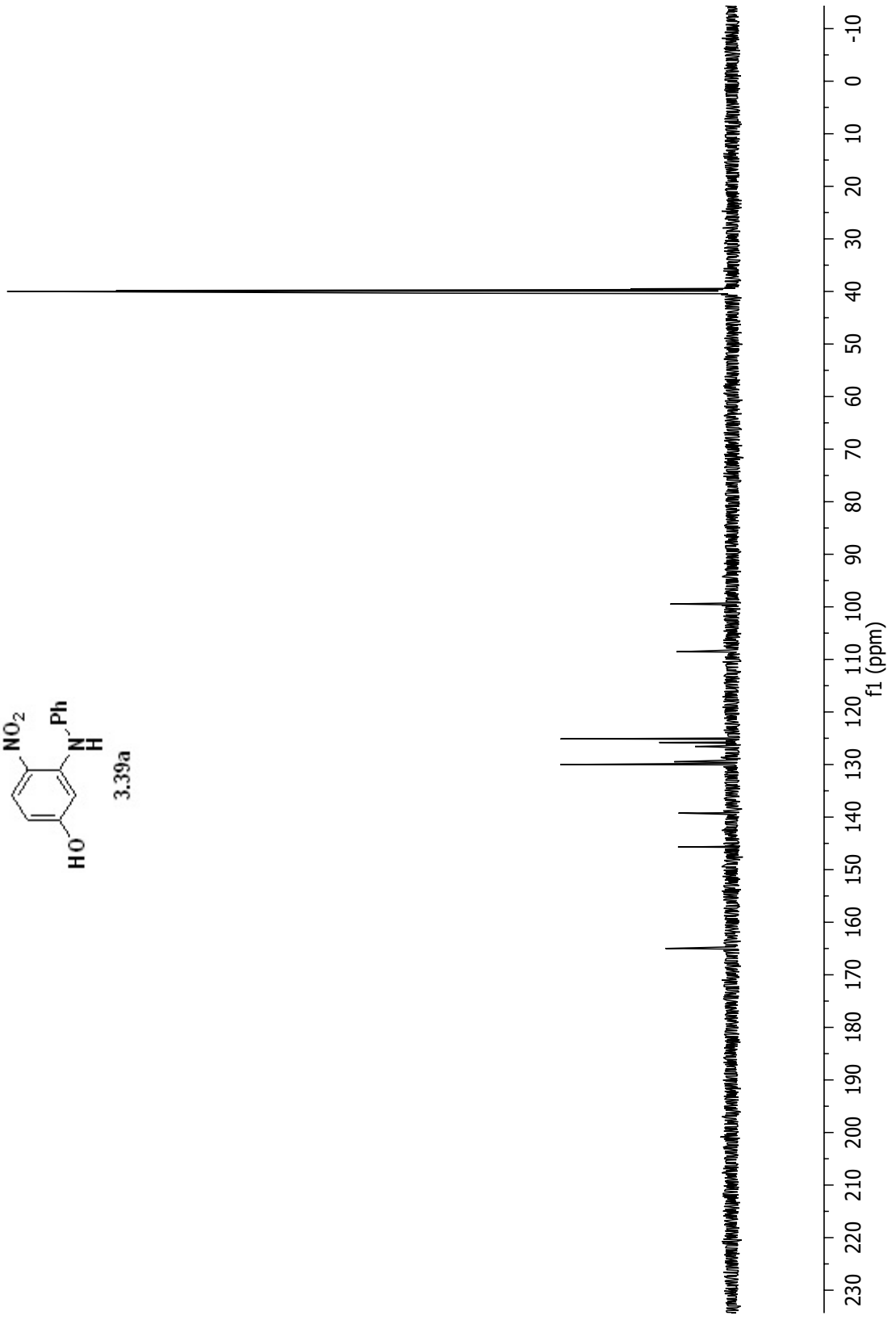

Cl
COOH
MeO
O
N
H
O
H
N
O
HOOC
O
N
N
H
O
3.15
1.00
0.34
0.36
0.29
0.91
0.85
0.60
0.70
0.43
0.41
0.91
0.47
1.06
0.78
0.53
0.52
13.0 12.5 12.0 11.5 11.0 10.5 10.0 9.5 9.0 8.5 8.0 7.5 7.0 6.5 6.0 5.5 5.0 4.5 4.0 3.5 3.0 2.5 2.0 1.5
f1 (ppm)

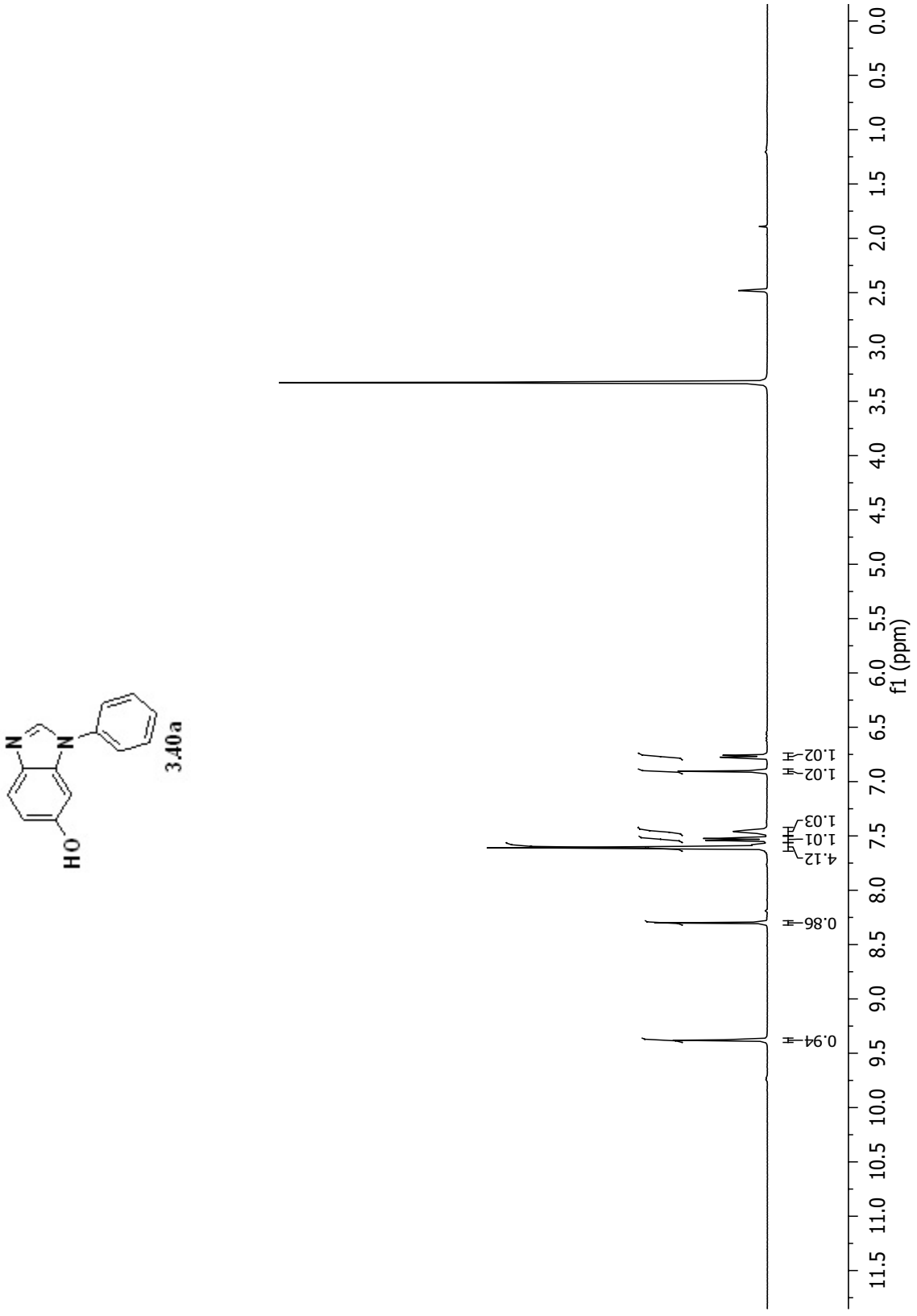

Cl
COOH
MeO
O
N
H
O
H
N
O
HOOC
O
N
N
H
O
3.15
200
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
f1 (ppm)

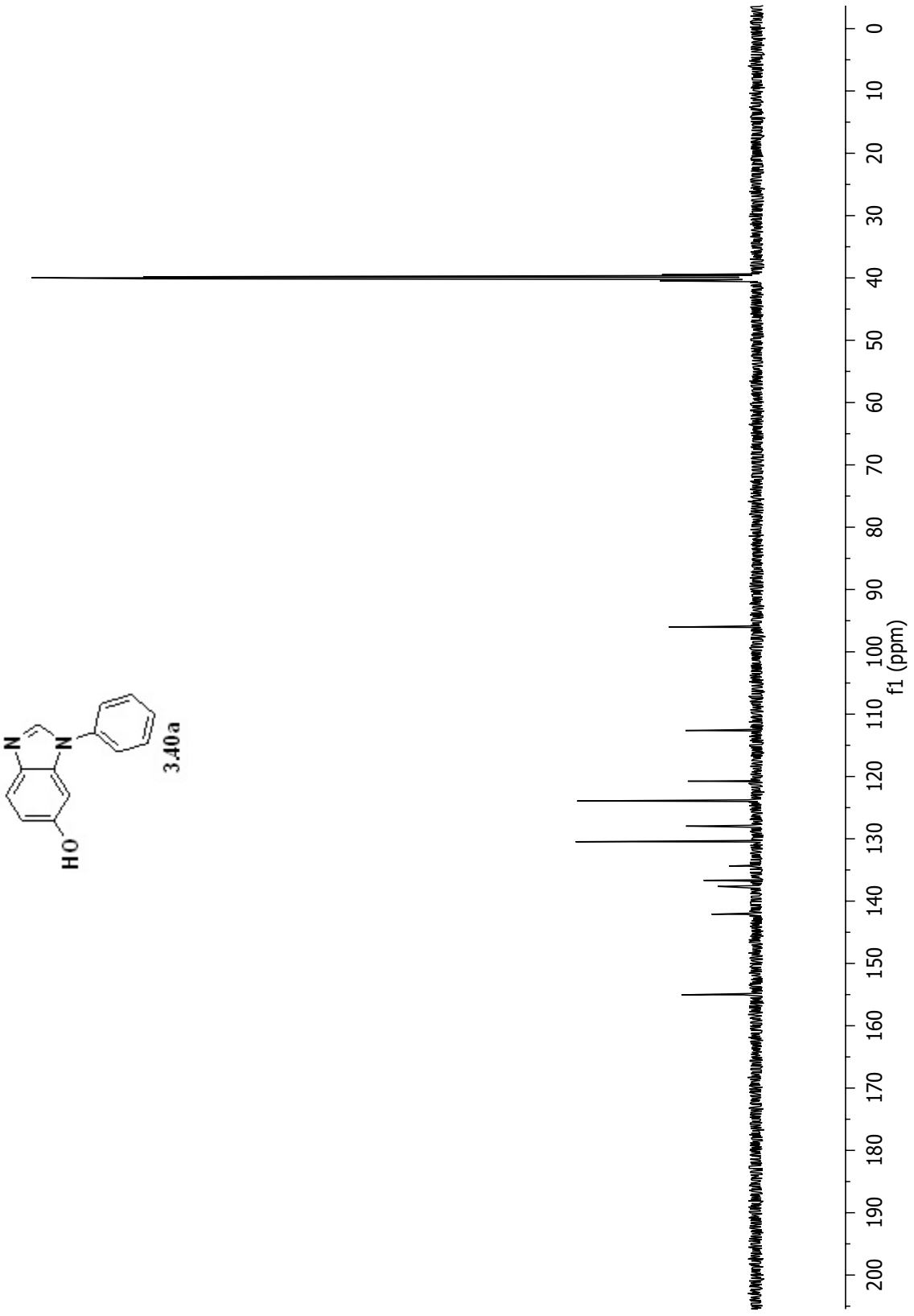


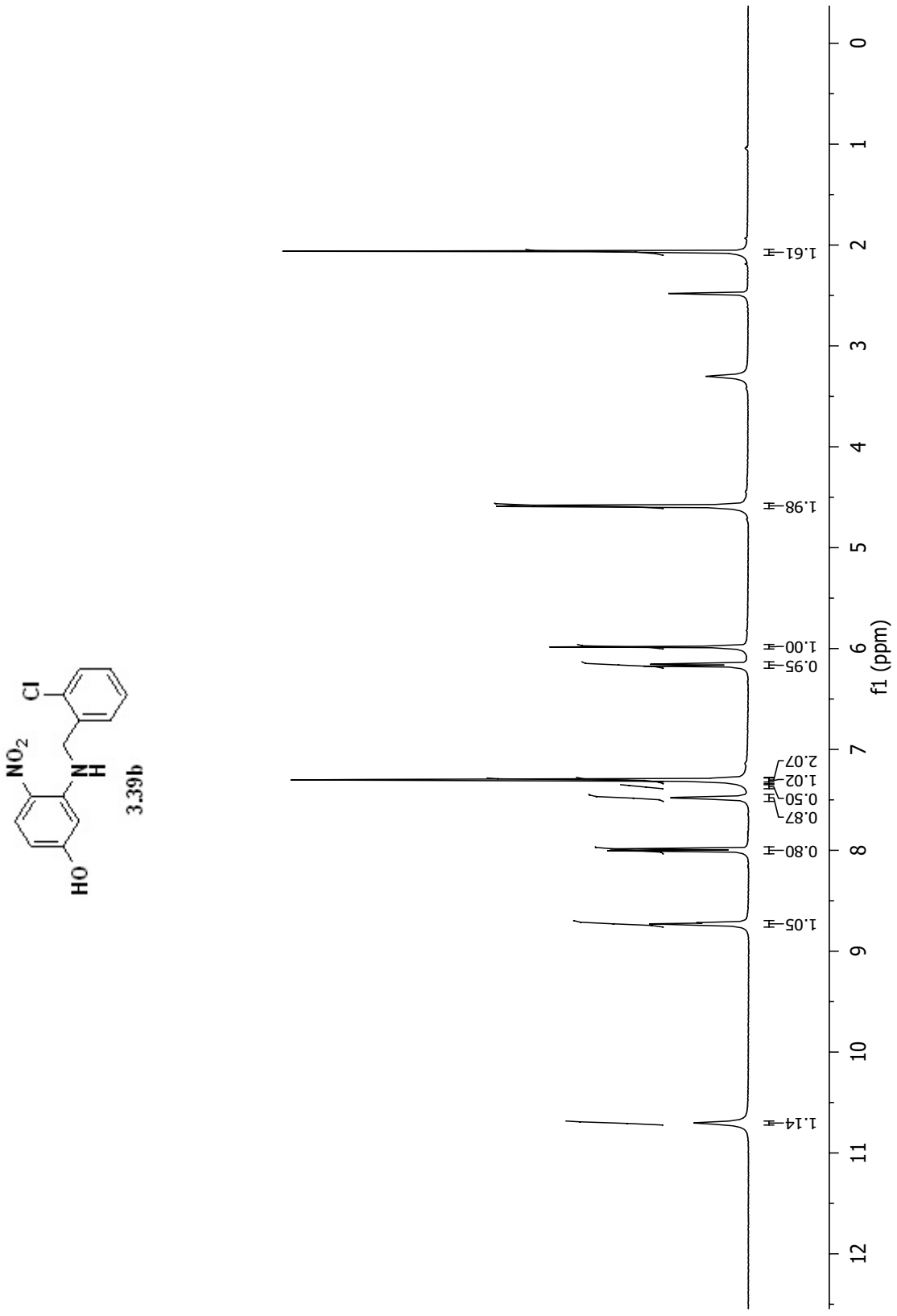
3.39a

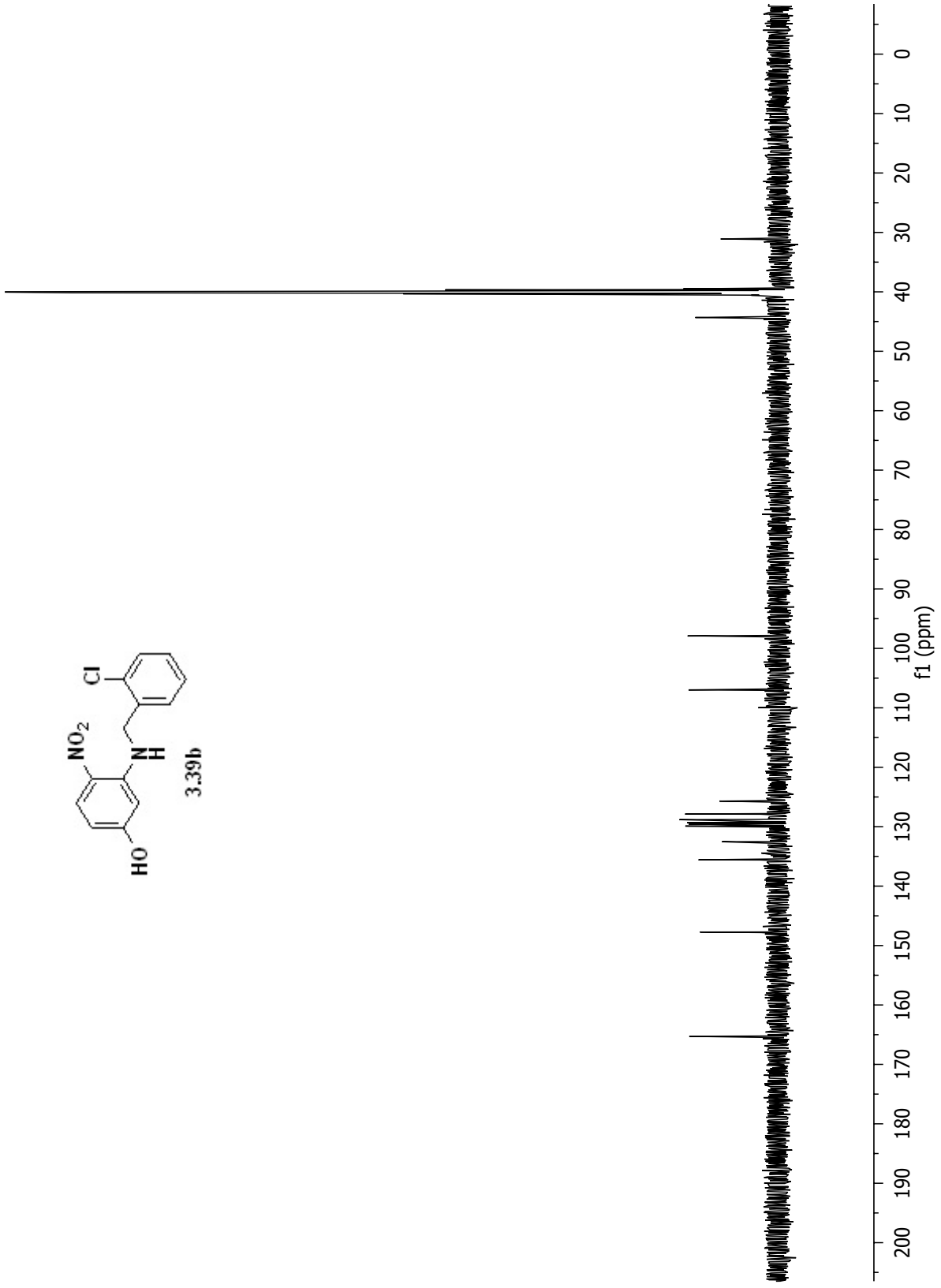


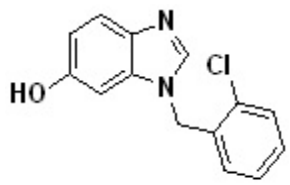

NO2
Ph
N
H
HO
3.39a
f1 (ppm)
-10
0
10
20
30
40
50
60
70
80
90
100
110
120
130
140
150
160
170
180
190
200
210
220
230


N
N
HO
3.40a
0.94
0.86
4.12
1.01
1.03
1.02
1.02
f1 (ppm)

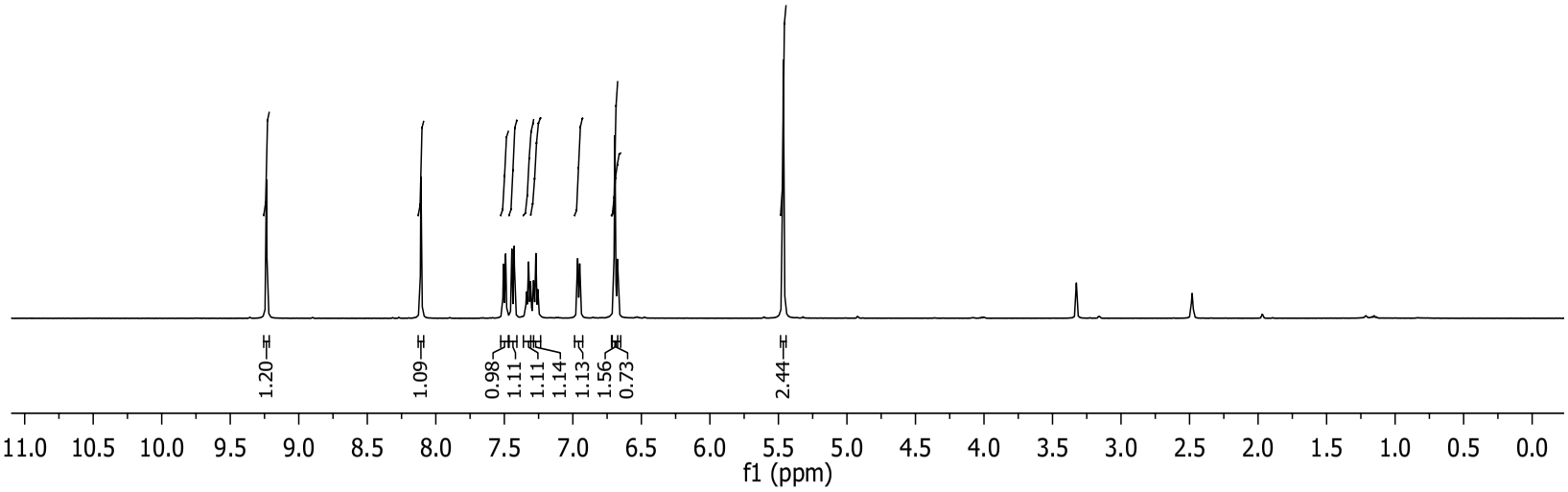

N
N
HO
3.40a
0
10
20
30
40
50
60
70
80
90
100
110
120
130
140
150
160
170
180
190
200
f1 (ppm)

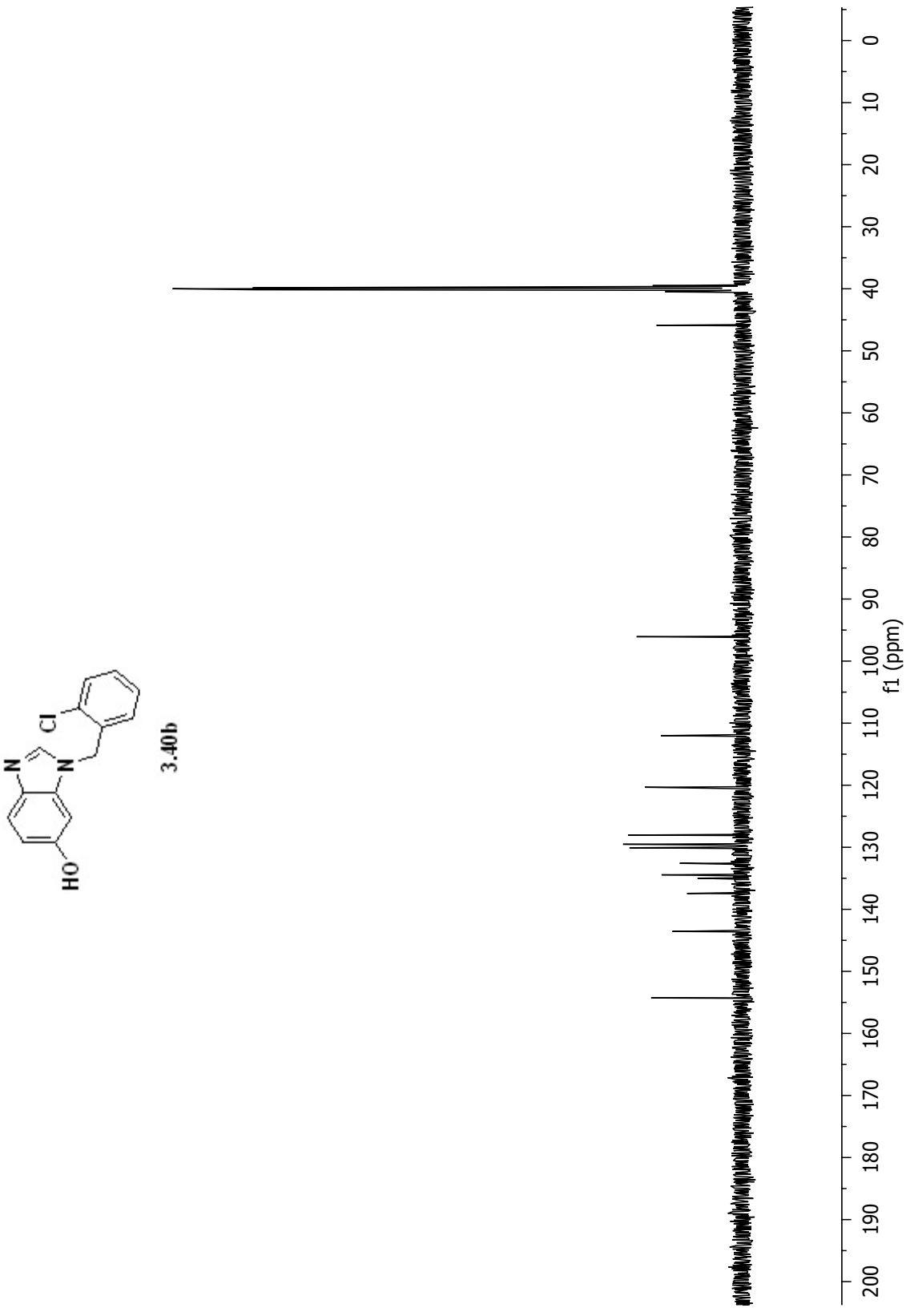

Cl
NO2
N
H
HO
3.39b
1.14
1.05
0.80
0.87
0.50
1.02
2.07
0.95
1.00
1.98
1.61
0
1
2
3
4
5
6
7
8
9
10
11
12
f1 (ppm)

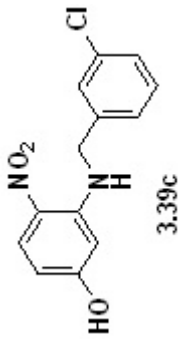

Cl
NO2
N
H
HO
3.39b
0
10
20
30
40
50
60
70
80
90
100
110
120
130
140
150
160
170
180
190
200
f1 (ppm)


N
Cl
HO
N

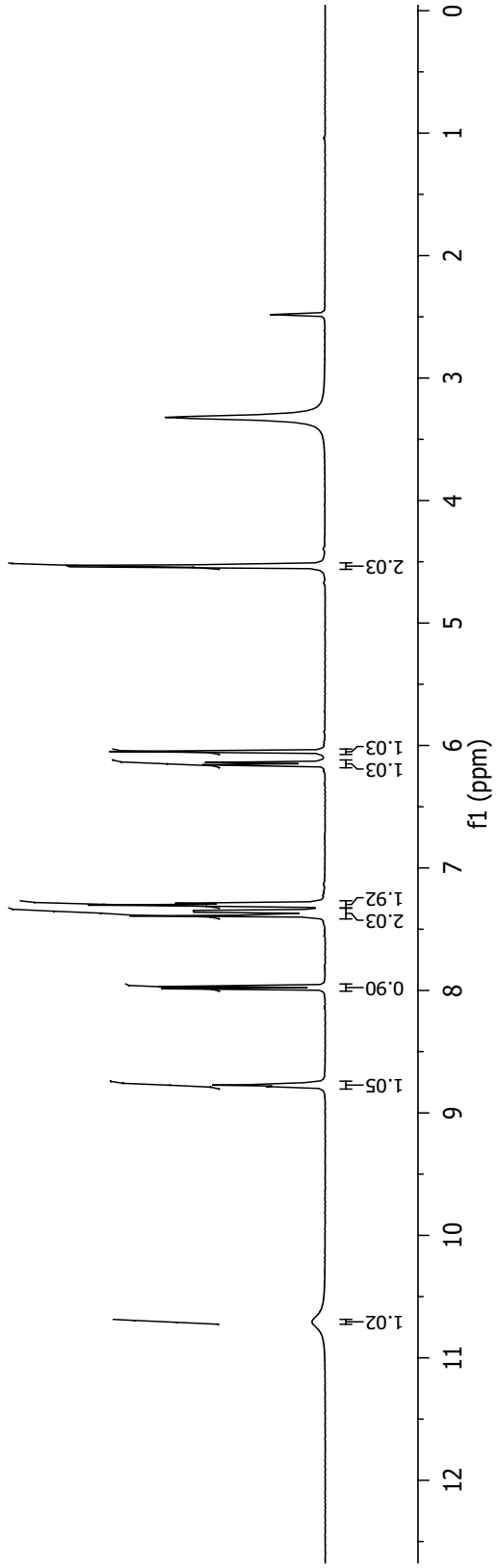
3.40b

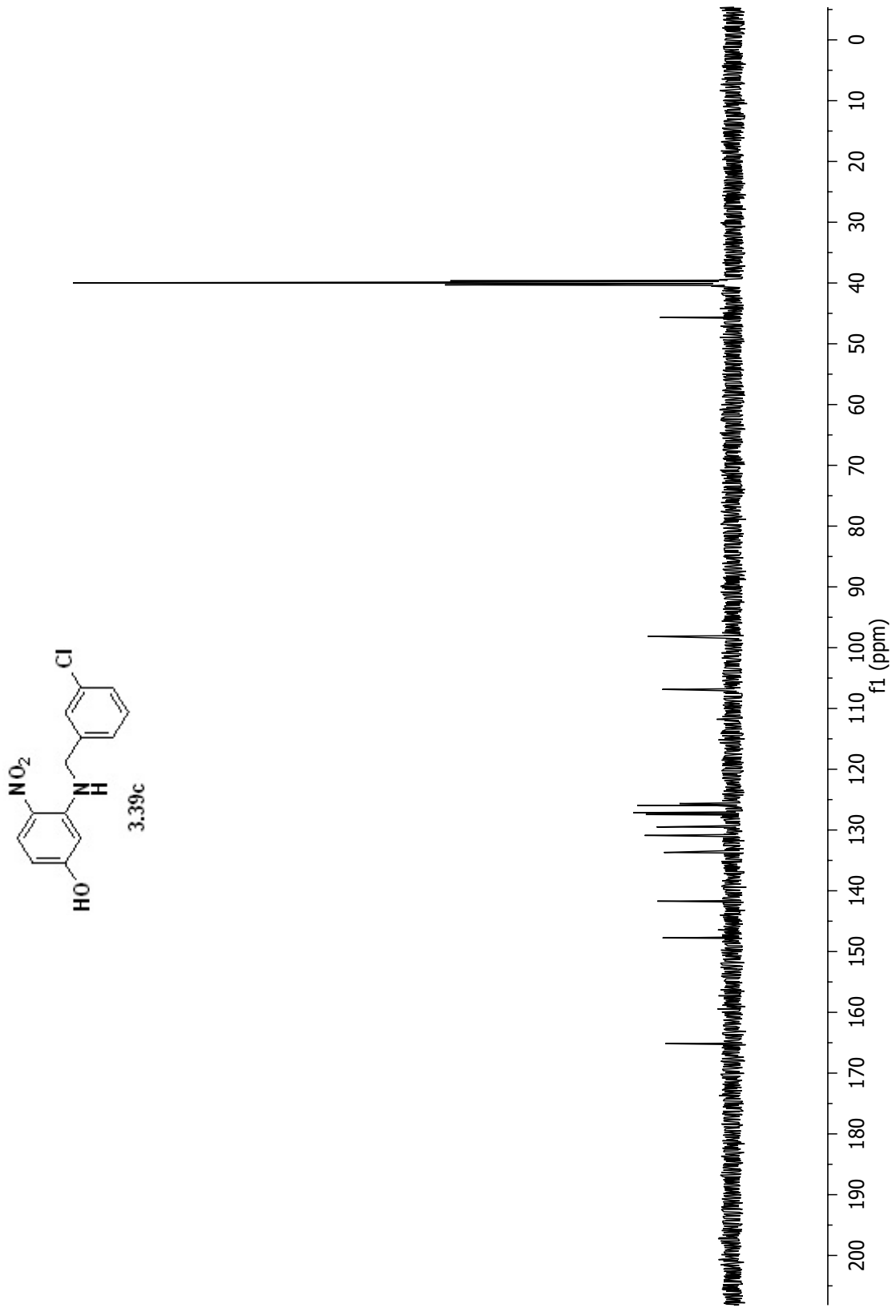

1.20
1.09
0.98
1.11
1.11
1.14
1.13
1.56
0.73
2.44
11.0 10.5 10.0 9.5 9.0 8.5 8.0 7.5 7.0 6.5 6.0 5.5 5.0 4.5 4.0 3.5 3.0 2.5 2.0 1.5 1.0 0.5 0.0
f1 (ppm)

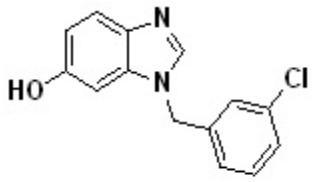

Cl
N
N
HO
3.40b
f1 (ppm)
0
10
20
30
40
50
60
70
80
90
100
110
120
130
140
150
160
170
180
190
200


Cl
NO2
N
H
HO

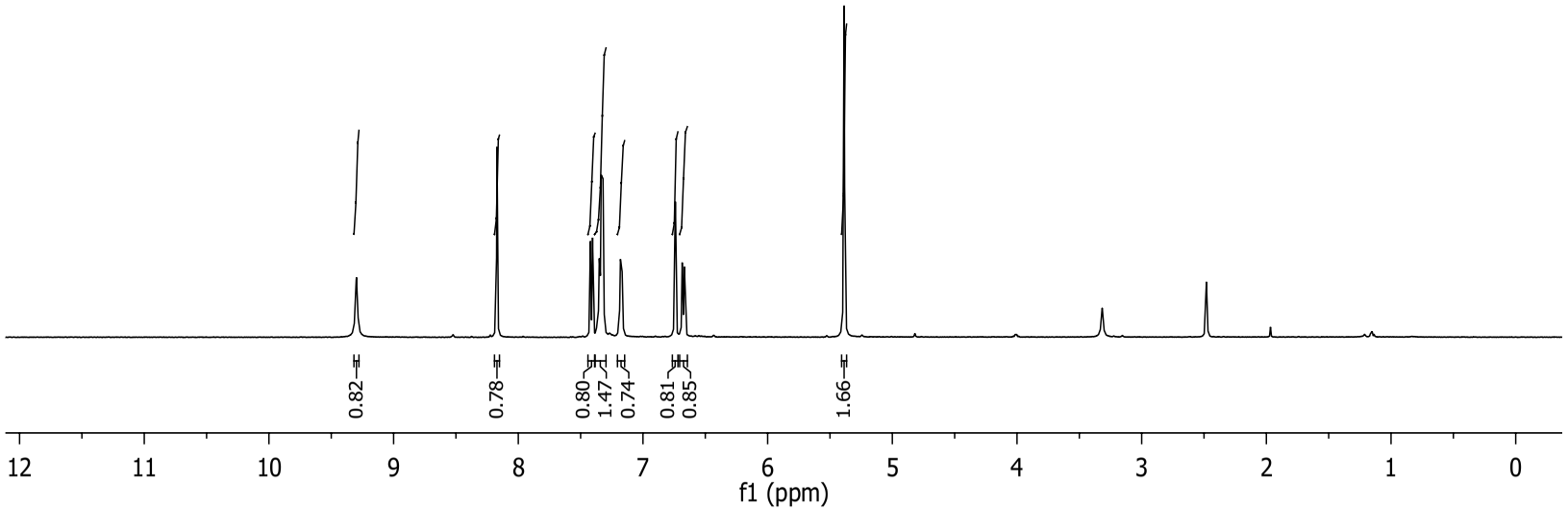
3.39c

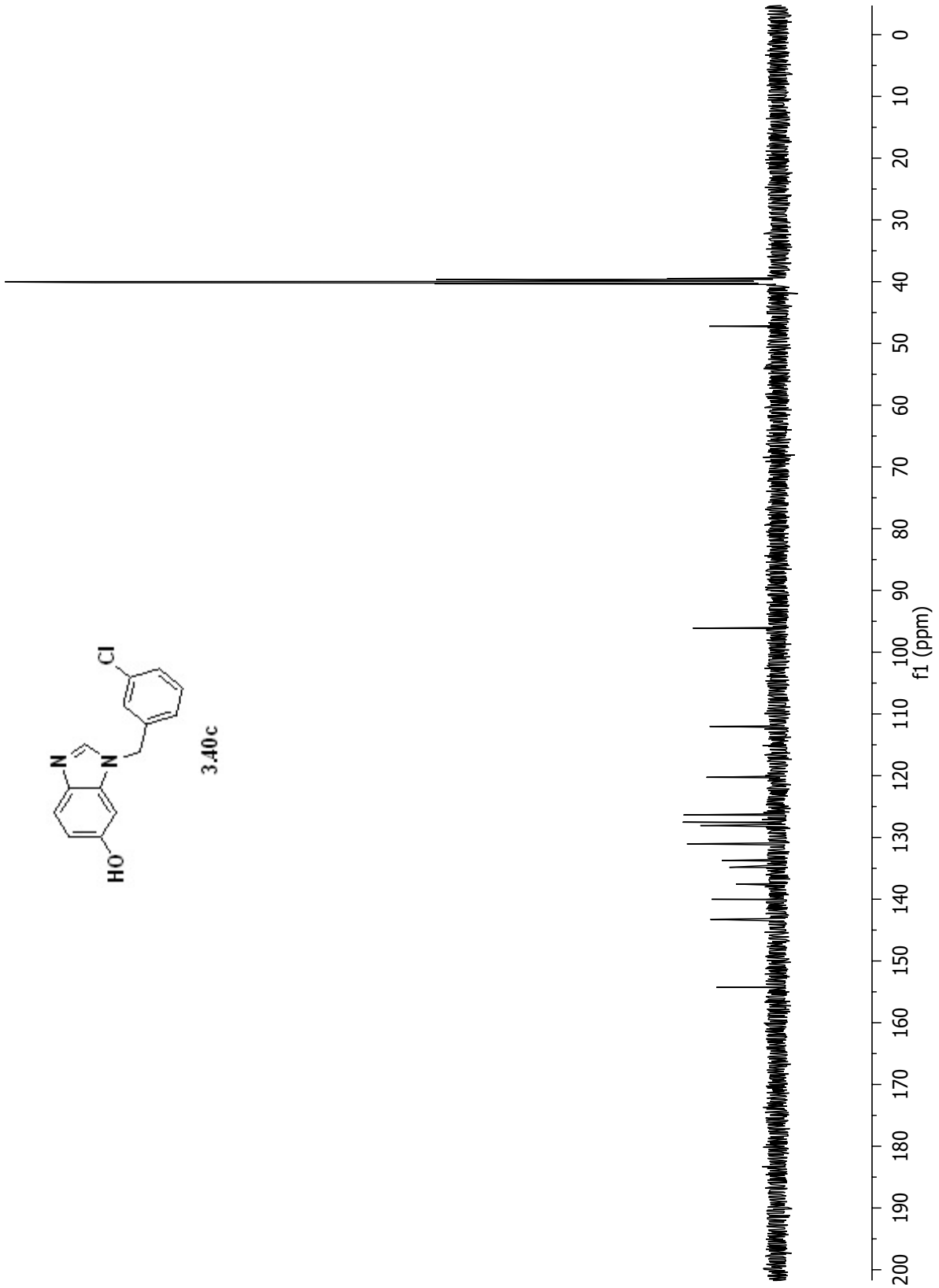

2.03
1.03
1.03
2.03
1.92
0.90
1.05
1.02
0
1
2
3
4
5
6
7
8
9
10
11
12
f1 (ppm)

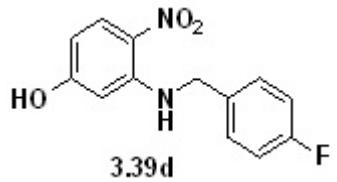

Cl
NO2
N
H
HO
3.39c
0
10
20
30
40
50
60
70
80
90
100
110
120
130
140
150
160
170
180
190
200
f1 (ppm)


N
HO
N
Cl

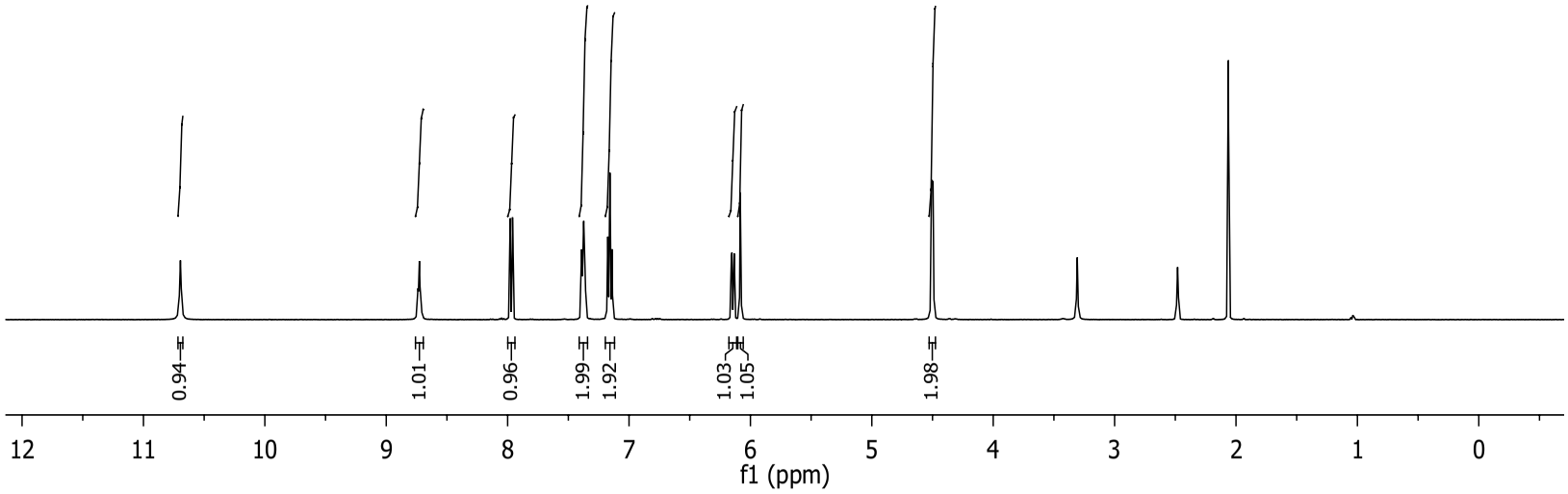
3.40c

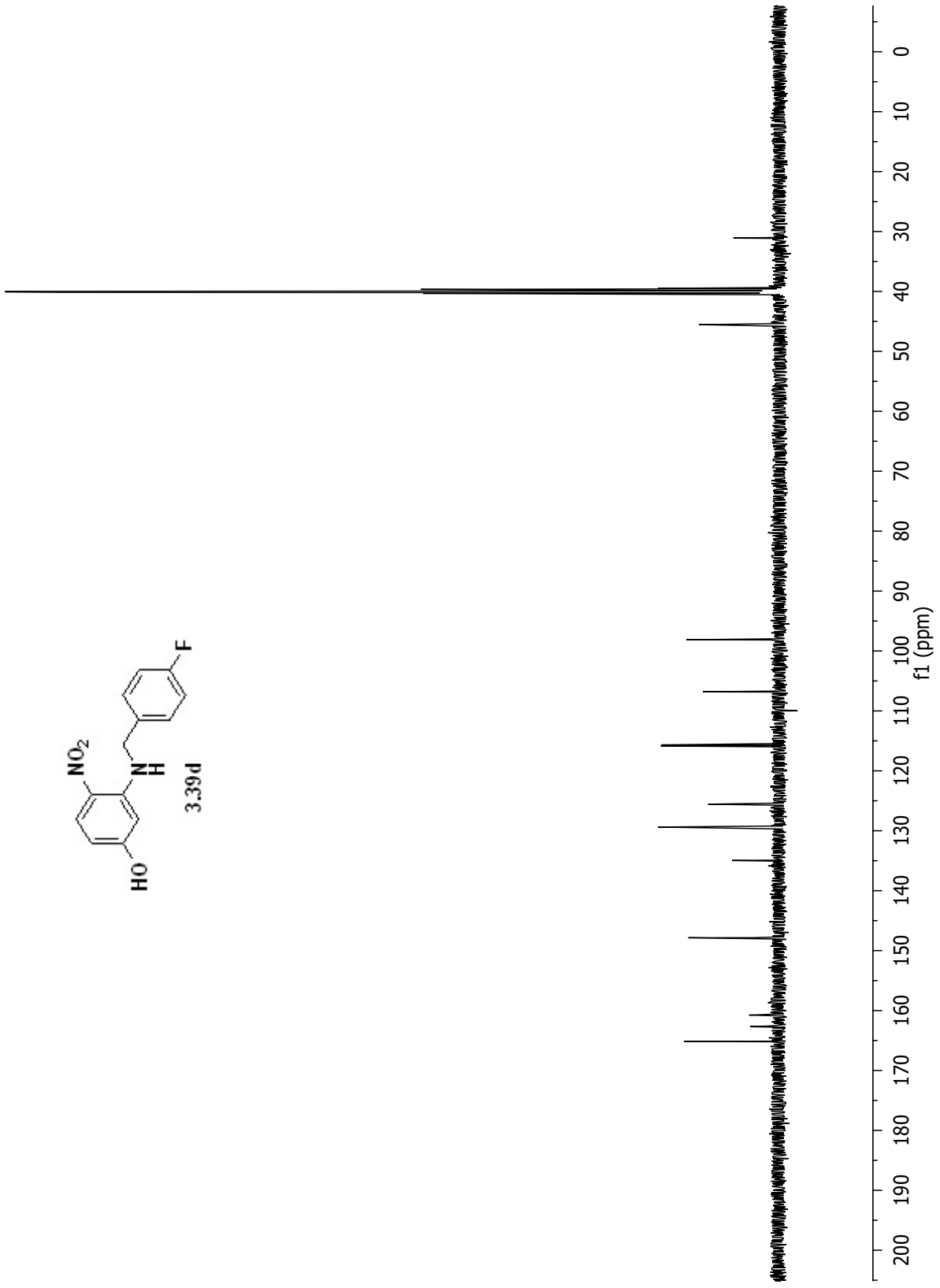

0.82
0.78
0.80
1.47
0.74
0.81
0.85
1.66
12
11
10
9
8
7
6
5
4
3
2
1
0
f1 (ppm)

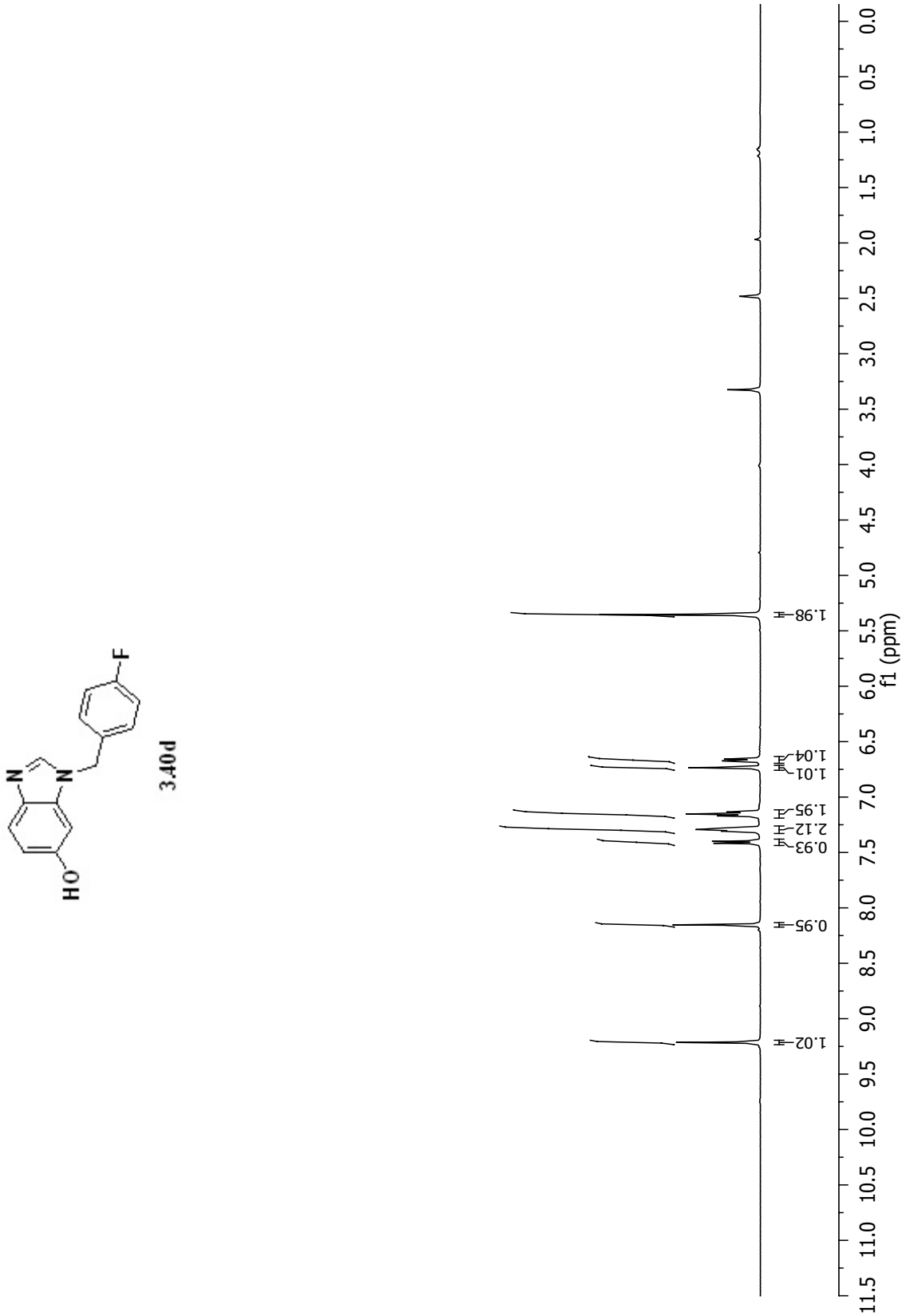

Cl
N
N
HO
3.40c
f1 (ppm)
0
10
20
30
40
50
60
70
80
90
100
110
120
130
140
150
160
170
180
190
200

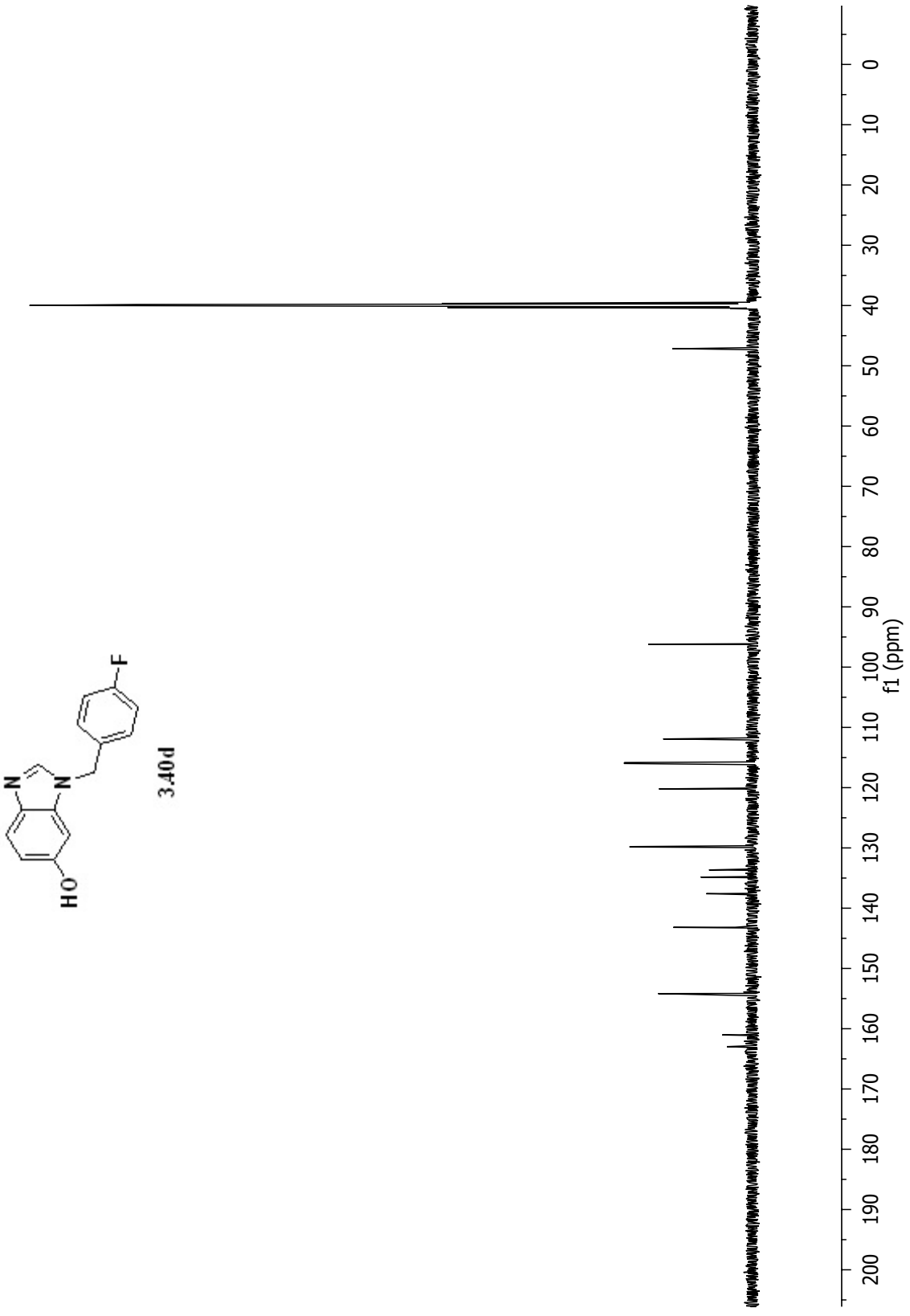

NO2
HO
N
H
F

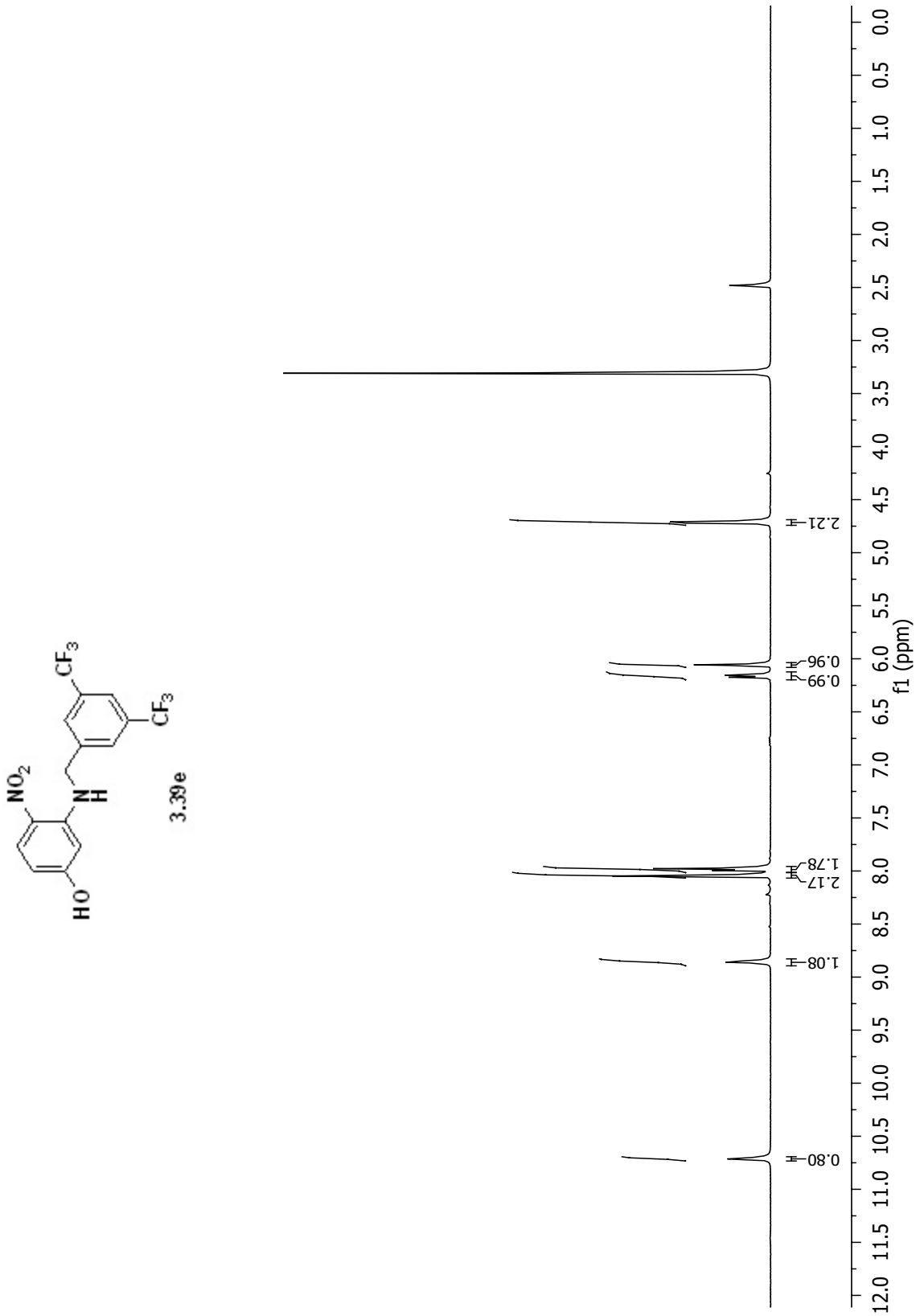
3.39d

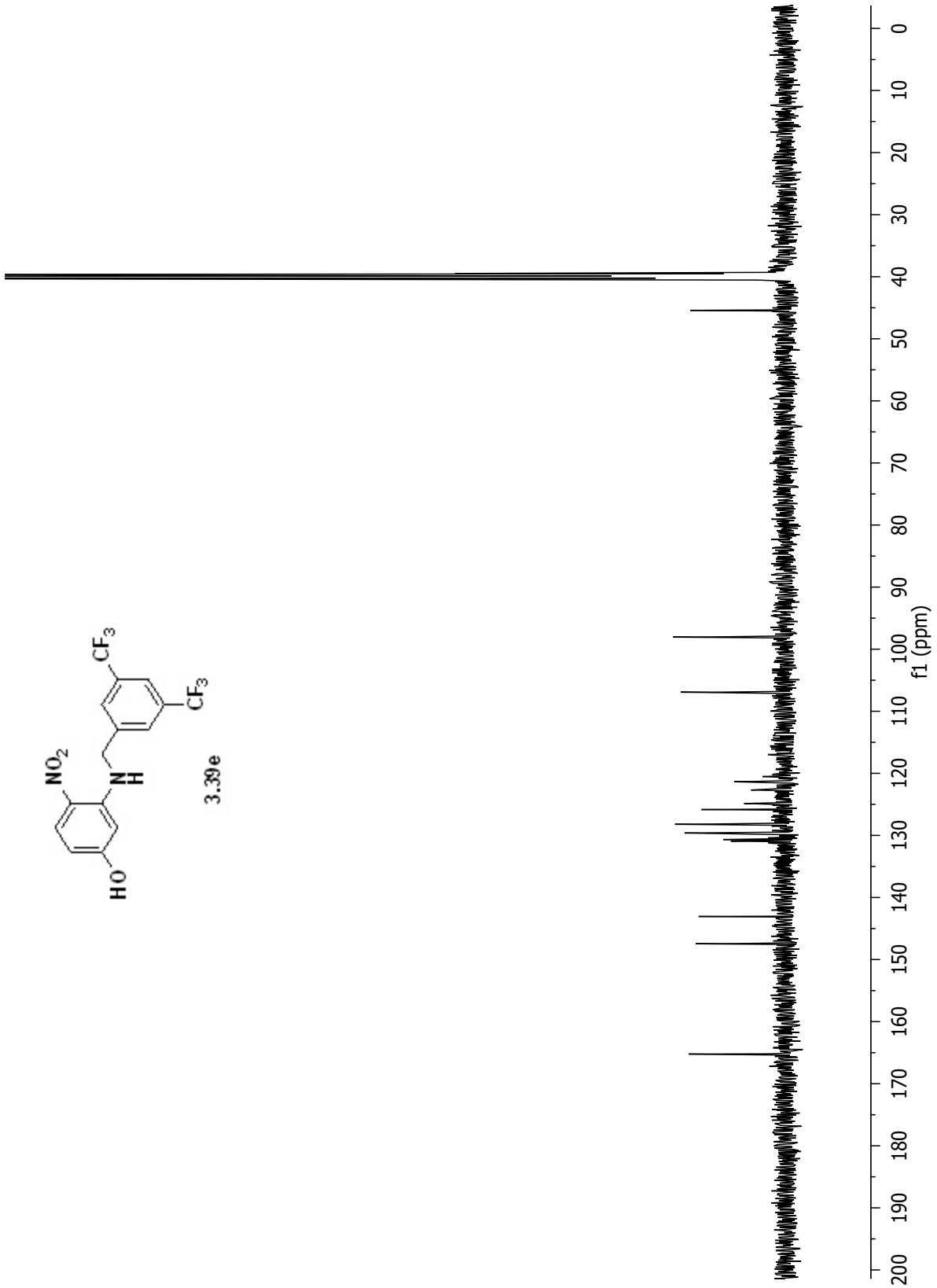

0.94
1.01
0.96
1.99
1.92
1.03
1.05
1.98
12
11
10
9
8
7
6
5
4
3
2
1
0
f1 (ppm)

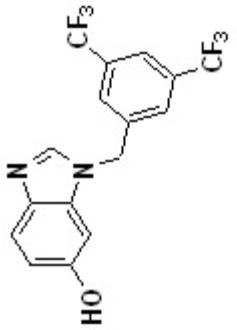

NO2
N
H
F
HO
3.39d
f1 (ppm)


F
N
N
HO
3.40d
1.98
1.01
1.04
1.95
2.12
0.93
0.95
1.02
f1 (ppm)

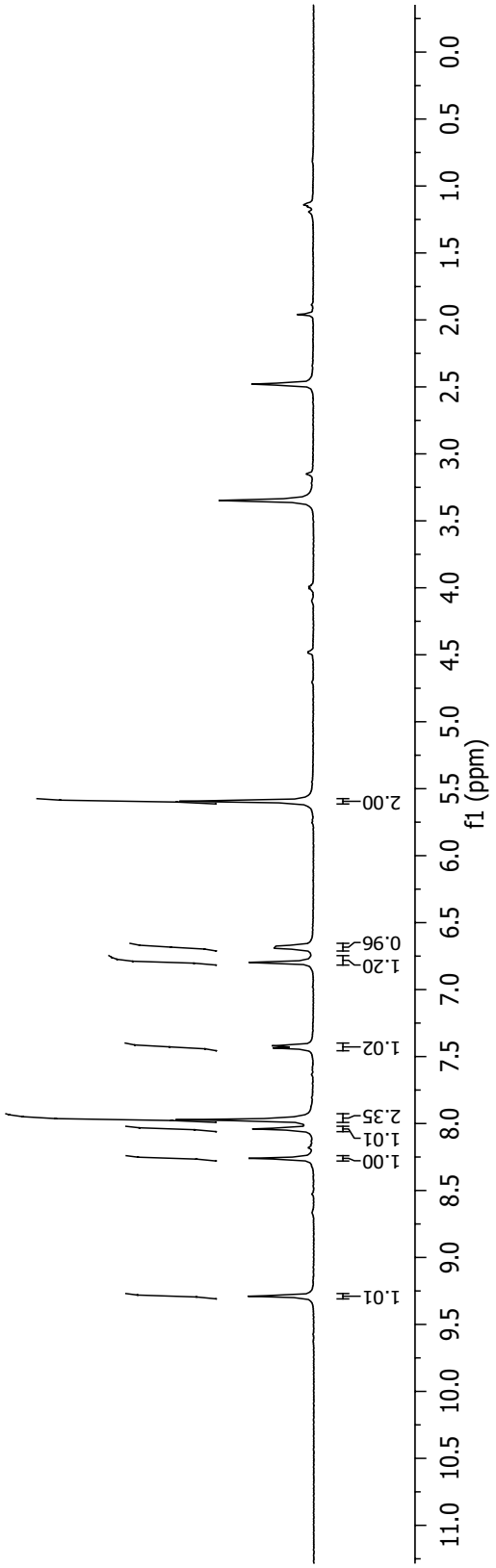

F
N
N
HO
3.40d
f1 (ppm)
0
10
20
30
40
50
60
70
80
90
100
110
120
130
140
150
160
170
180
190
200

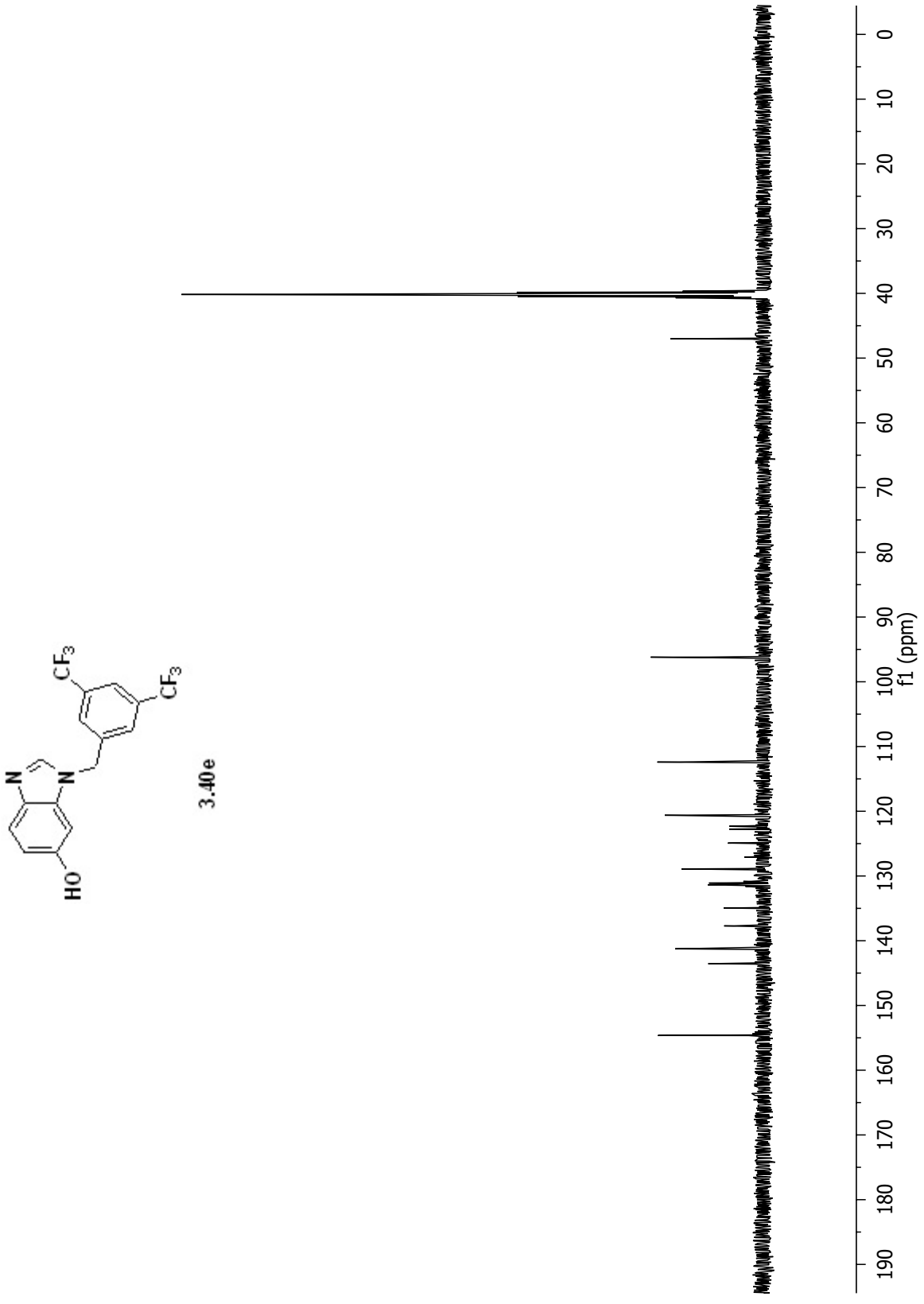

NO2
N
H
HO
CF3
CF3
3.39e
2.21
0.96
0.99
2.17
1.78
1.08
0.80
f1 (ppm)

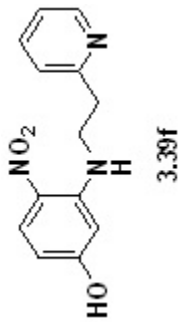

NO2
N
H
HO
CF3
CF3
3.39e
f1 (ppm)
0
10
20
30
40
50
60
70
80
90
100
110
120
130
140
150
160
170
180
190
200



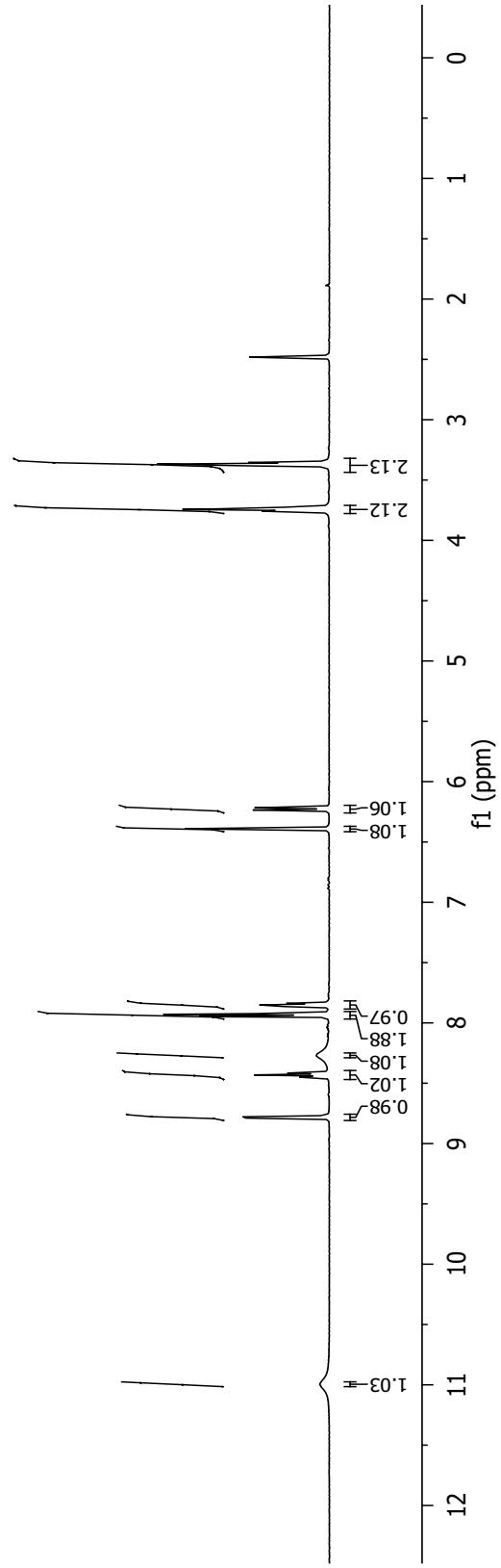
3.40e

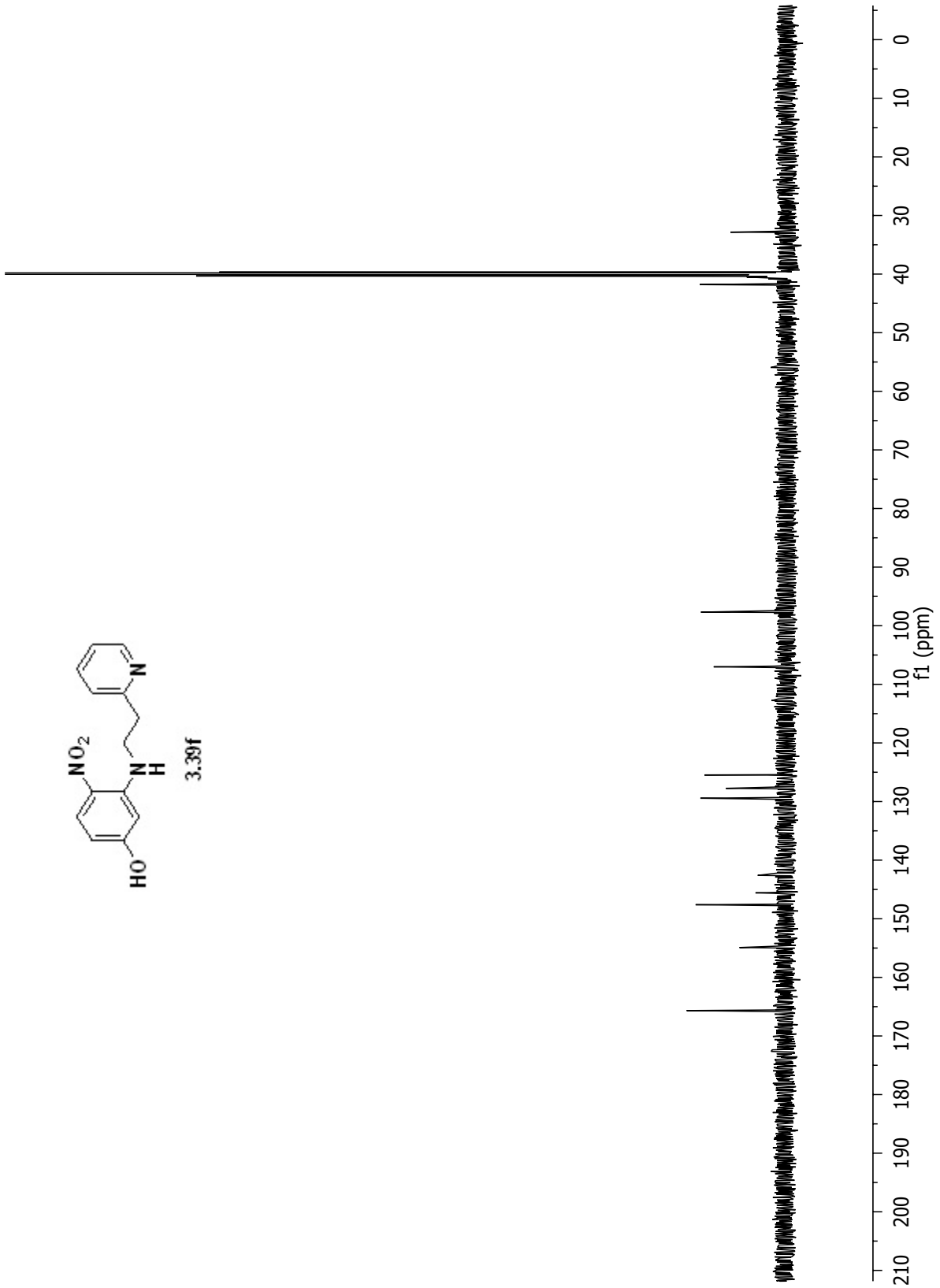


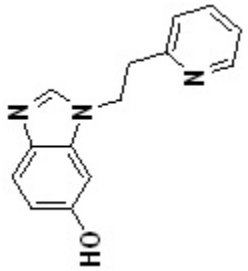

CF3
CF3
N
N
HO
3.40e
f1 (ppm)
0
10
20
30
40
50
60
70
80
90
100
110
120
130
140
150
160
170
180
190



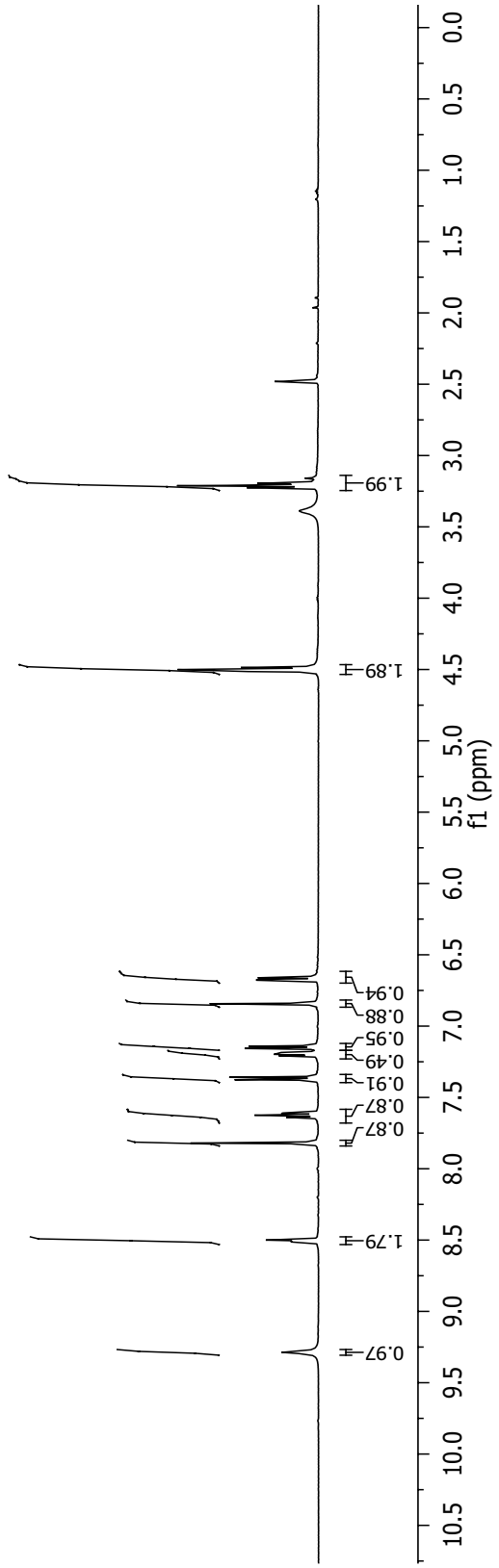
3.39f

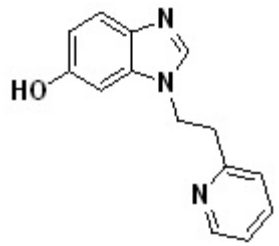



NO2
N
H
HO
N
3.39f
f1 (ppm)
0
10
20
30
40
50
60
70
80
90
100
110
120
130
140
150
160
170
180
190
200
210

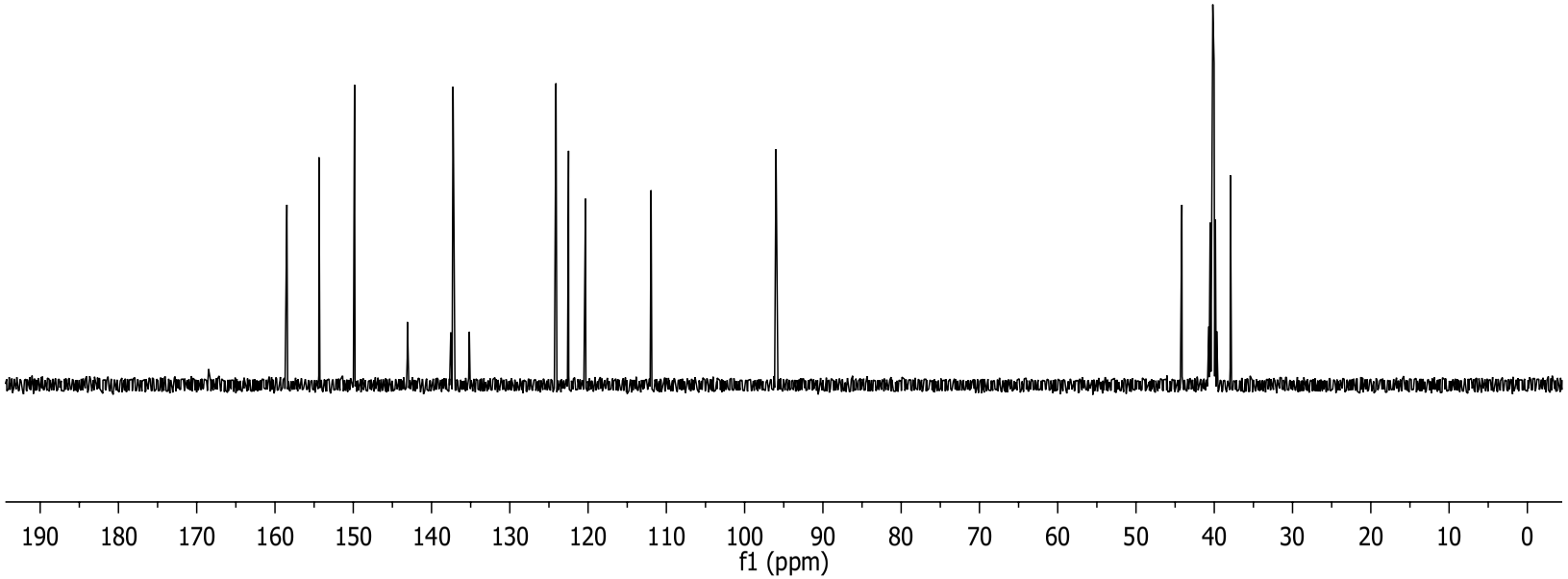

N
N
N
HO

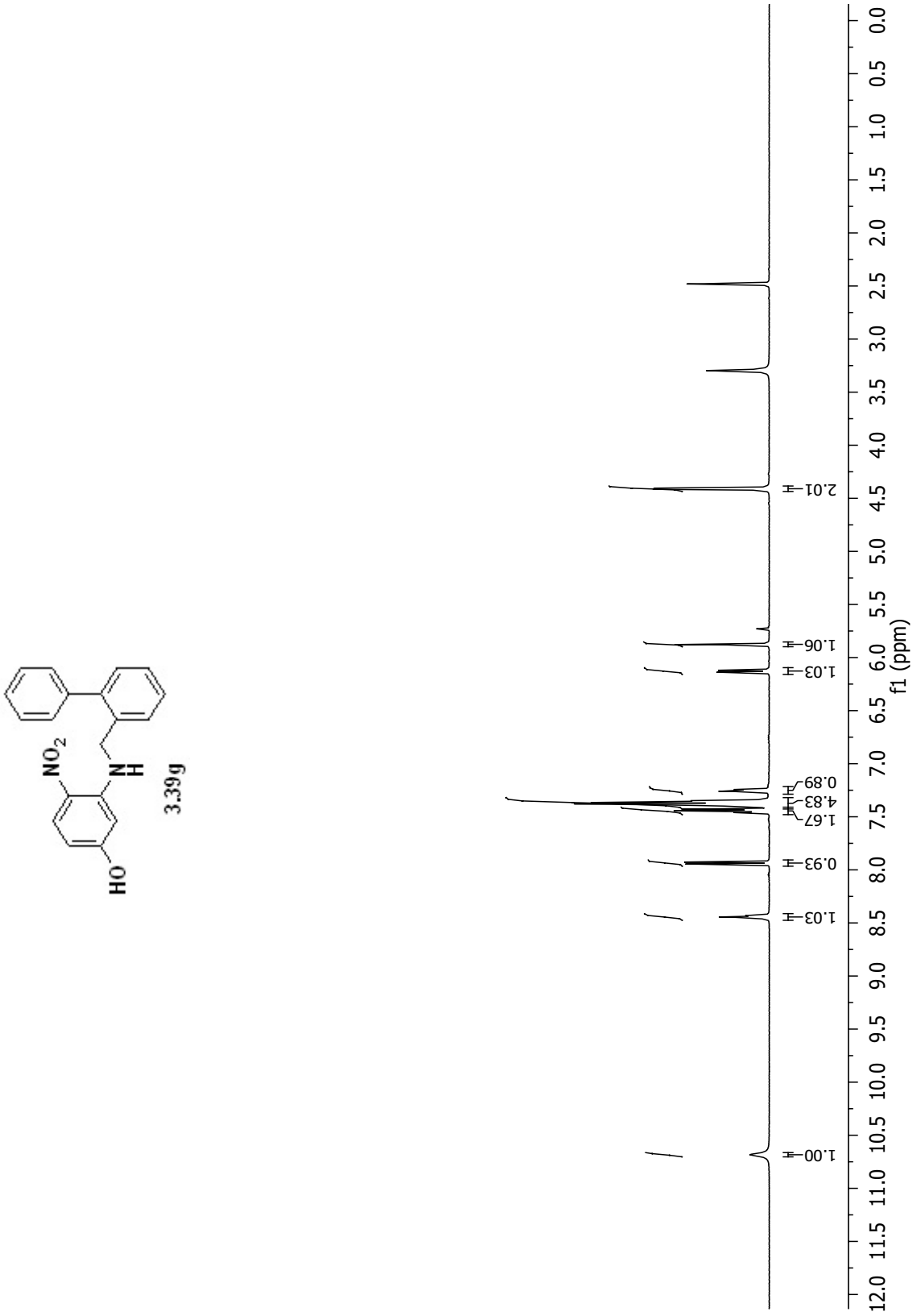
3.40f

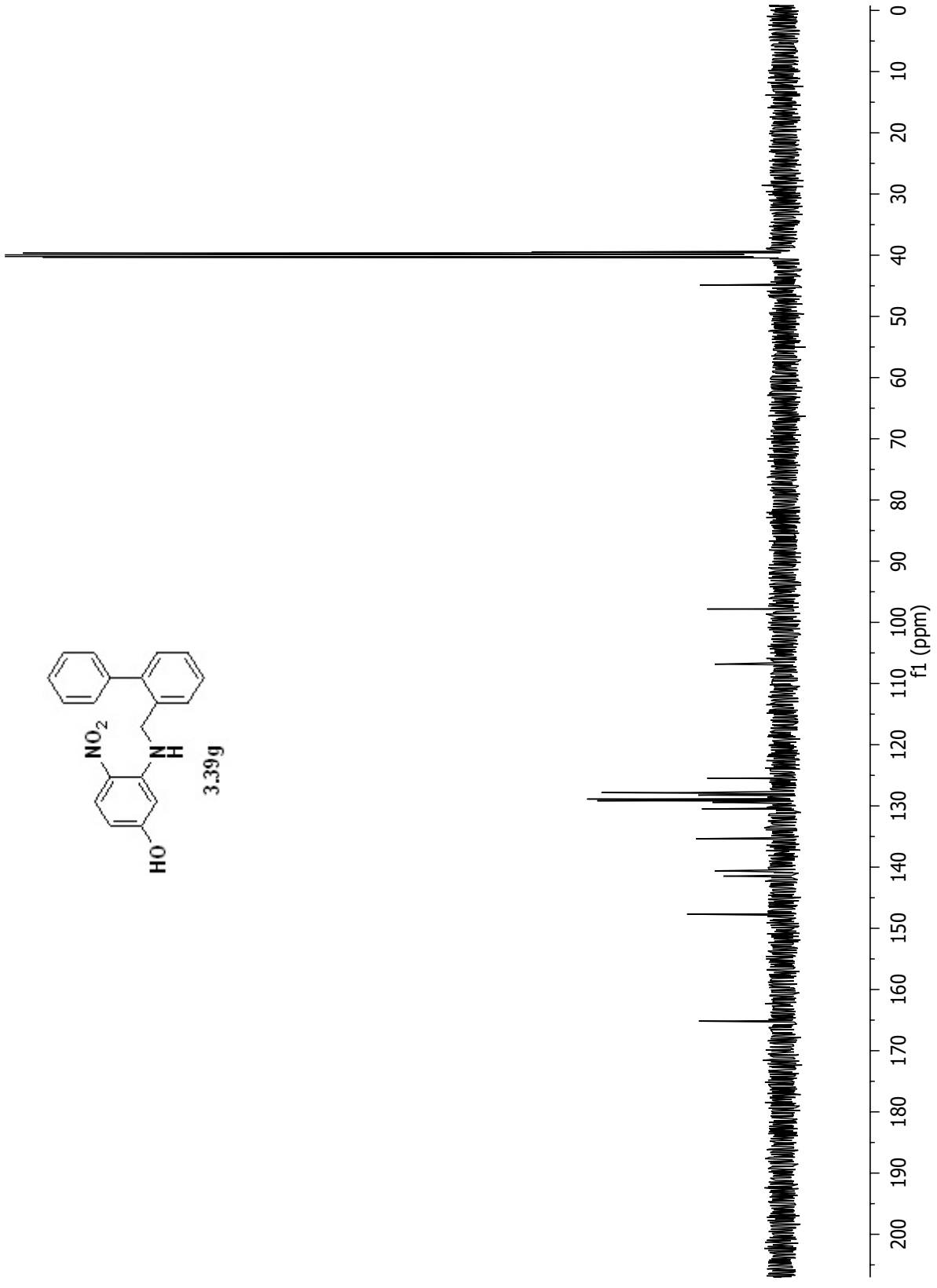

0.97
1.79
0.87
0.87
0.91
0.49
0.95
0.88
0.94
1.89
1.99
10.5
10.0
9.5
9.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
f1 (ppm)

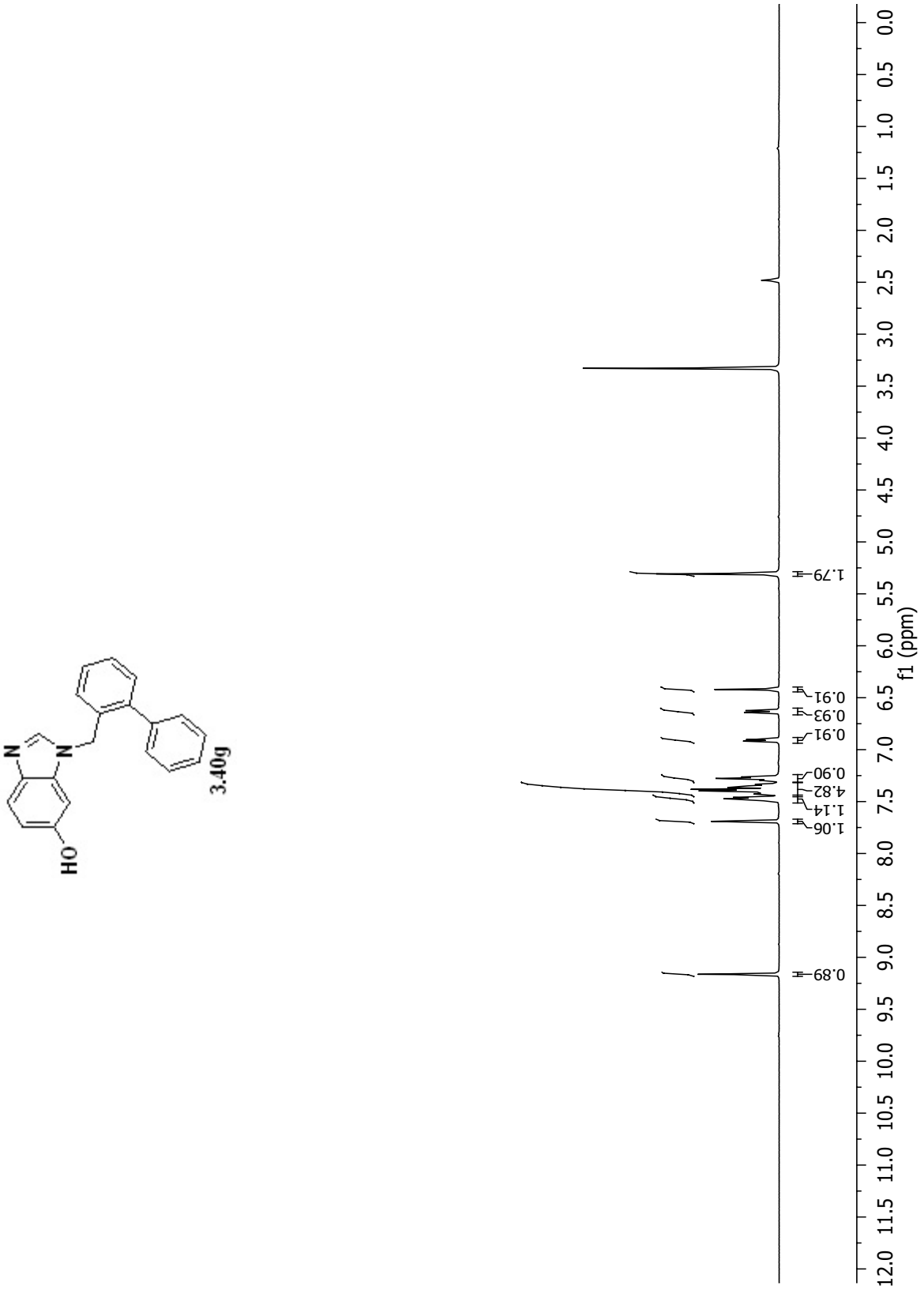

HO
N
N
N

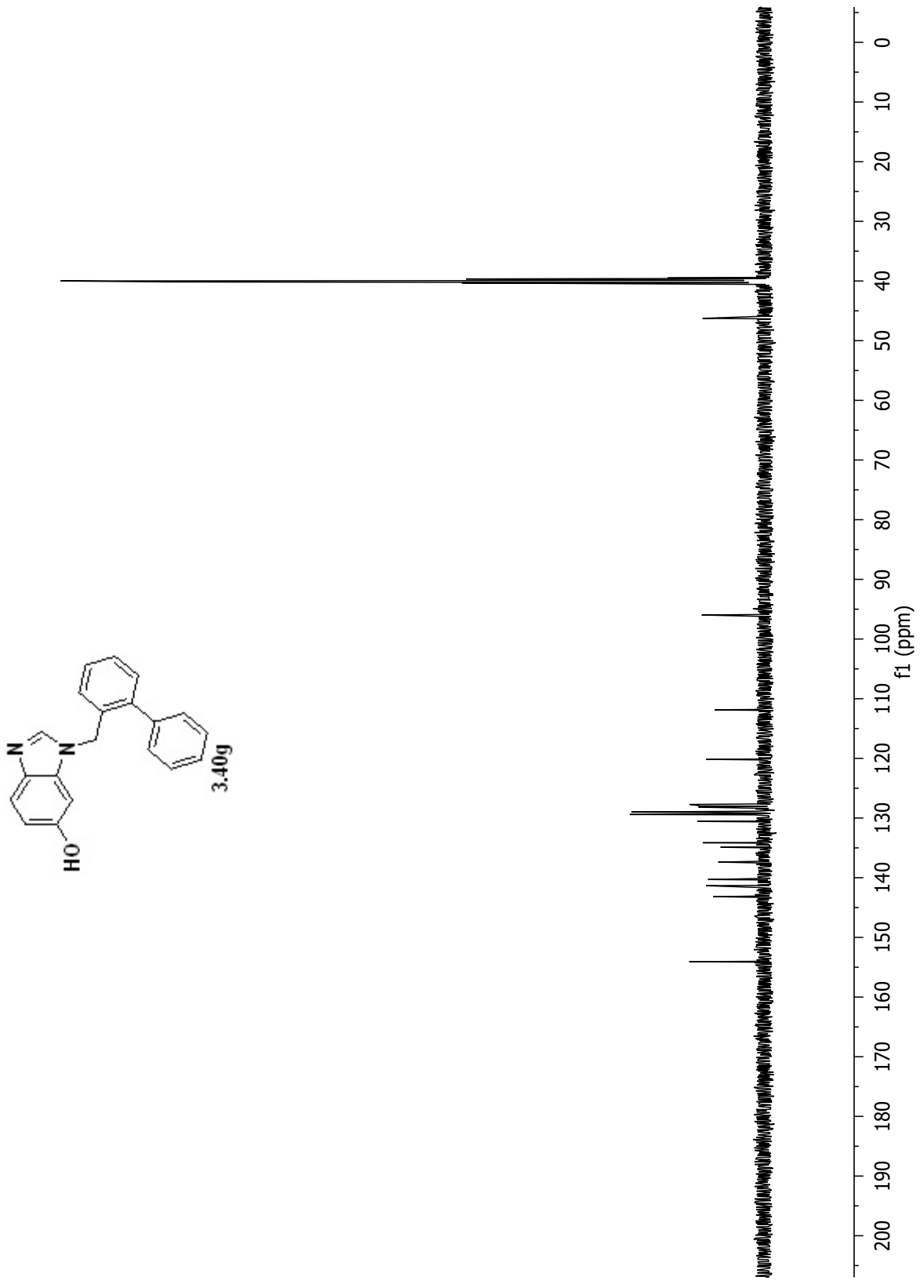
3.40f

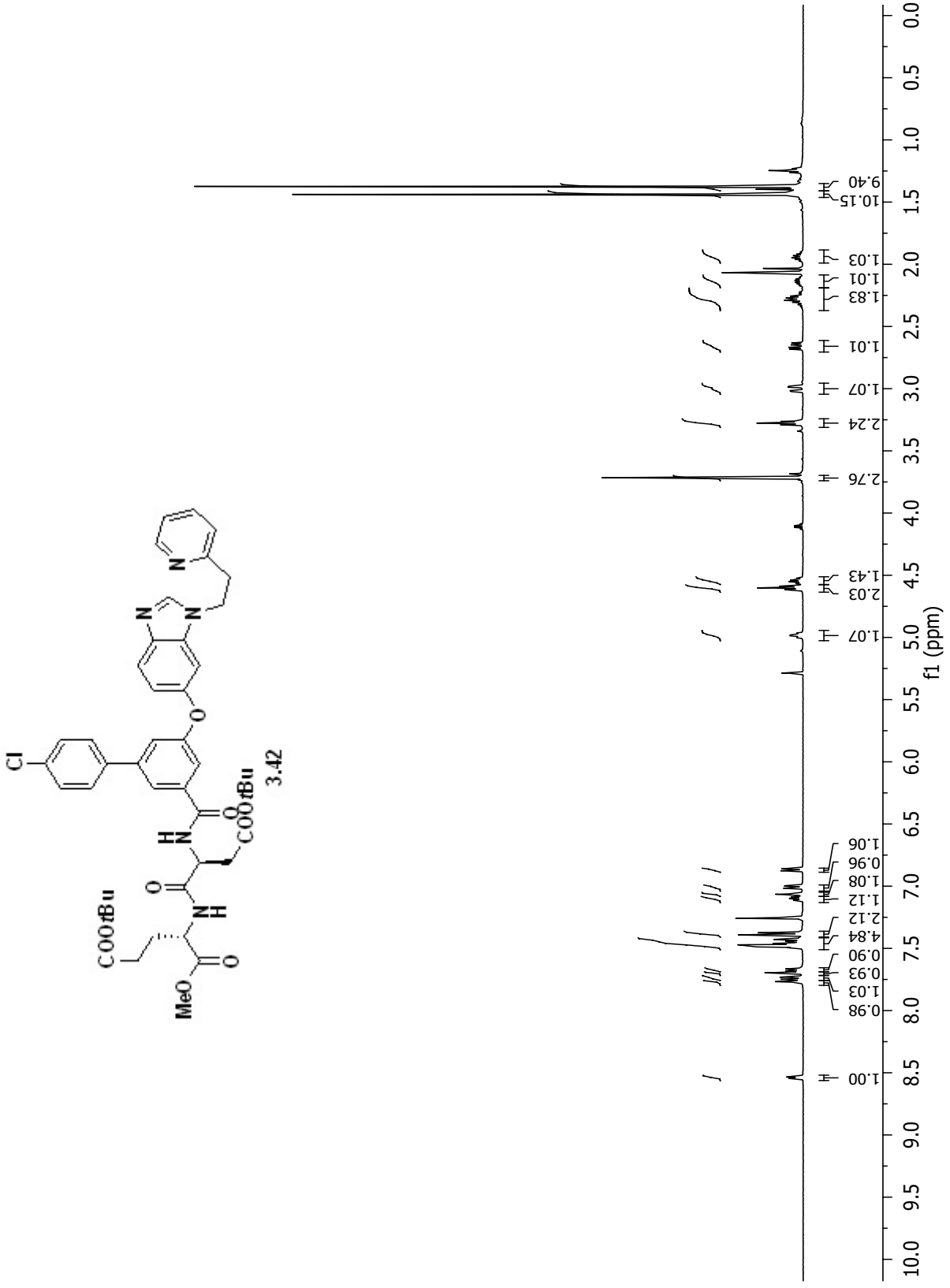

190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
f1 (ppm)

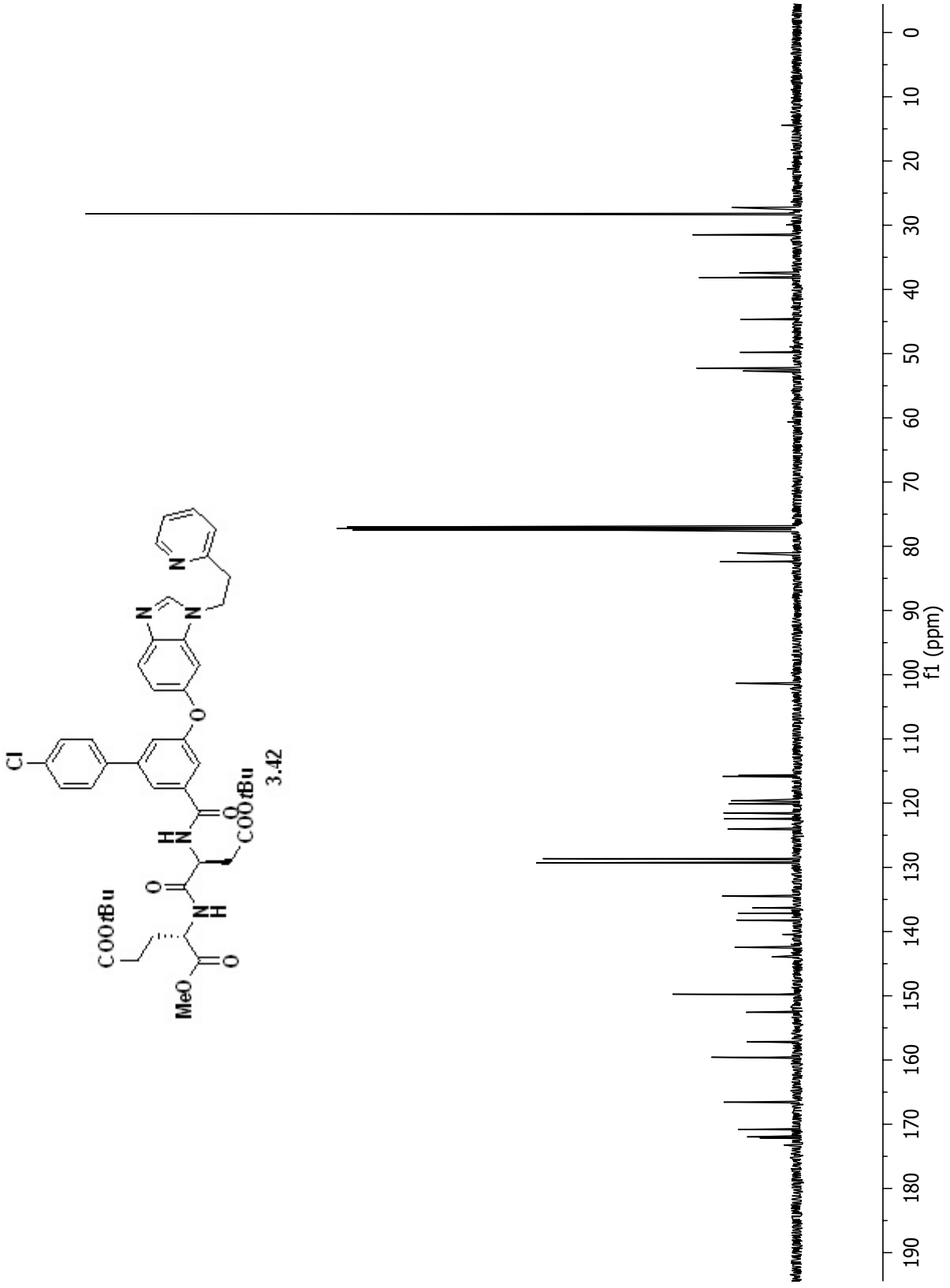

NO2
N
H
HO
3.39g
1.00
1.03
0.93
1.03
1.06
2.01
0.89
4.83
1.67
f1 (ppm)

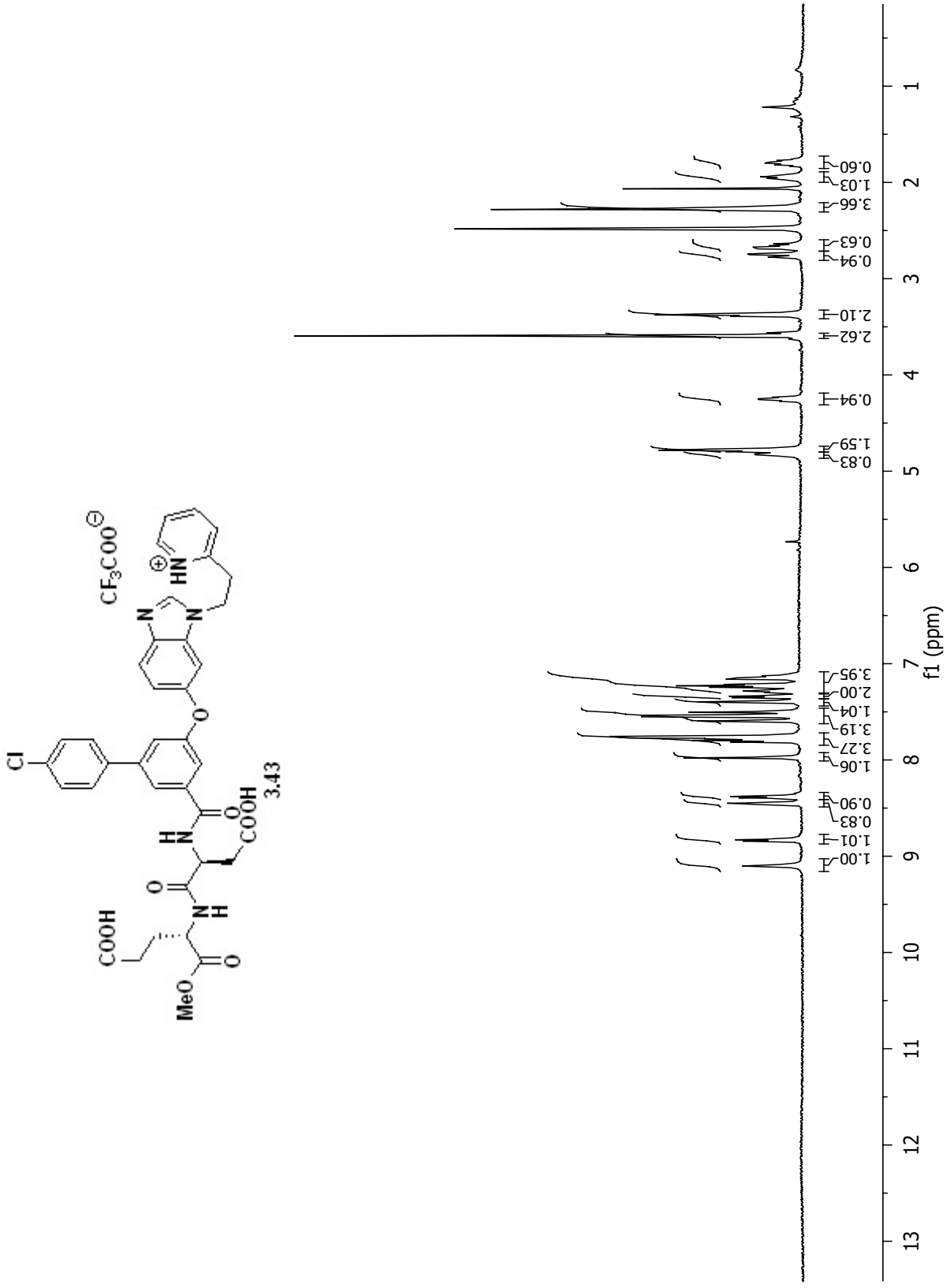

NO2
N
H
HO
3.39g
0
10
20
30
40
50
60
70
80
90
100
110
120
130
140
150
160
170
180
190
200
f1 (ppm)

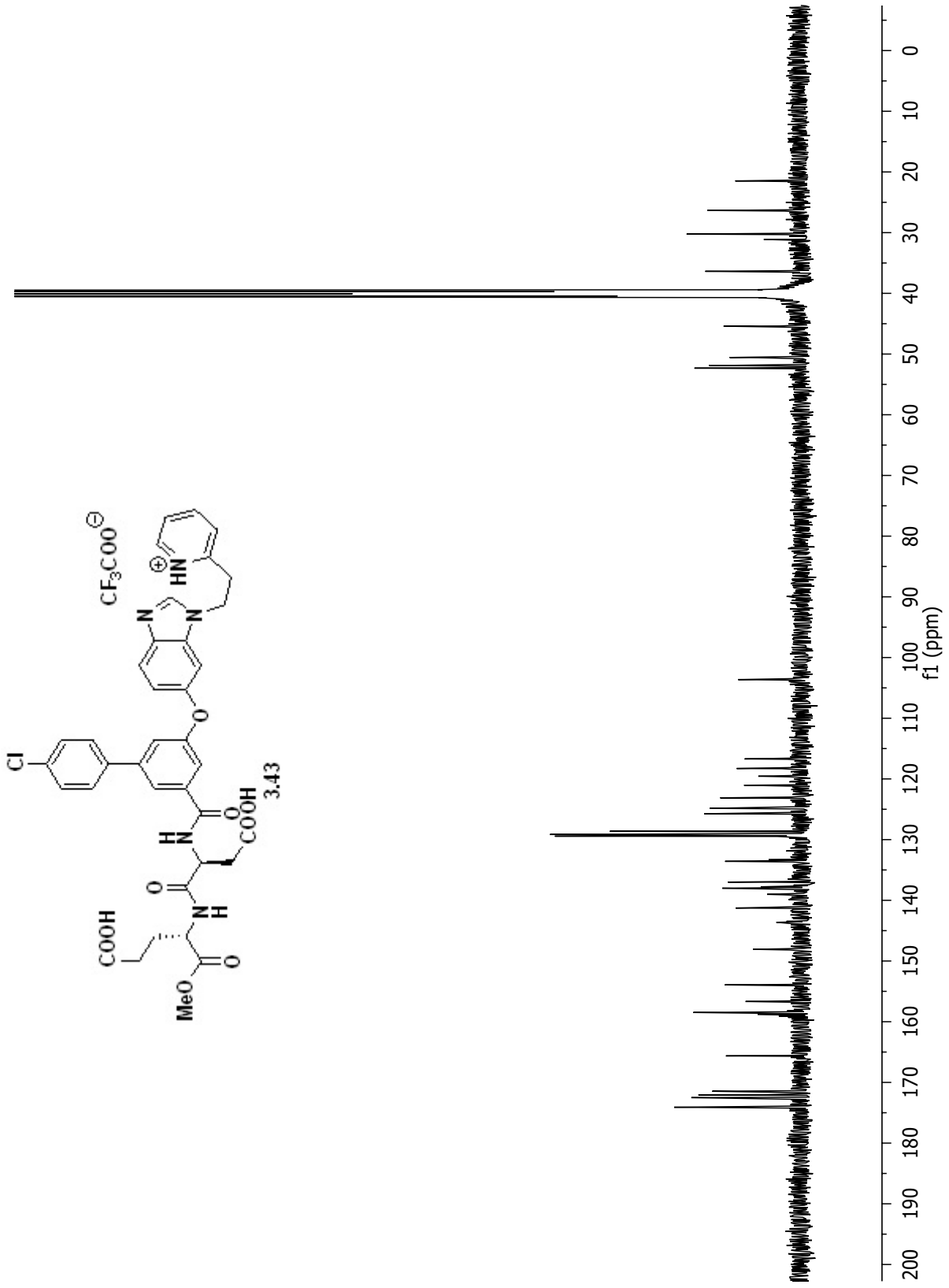

N
N
HO
3.40g
0.89
1.06
1.14
4.82
0.90
0.91
0.93
0.91
1.79
f1 (ppm)

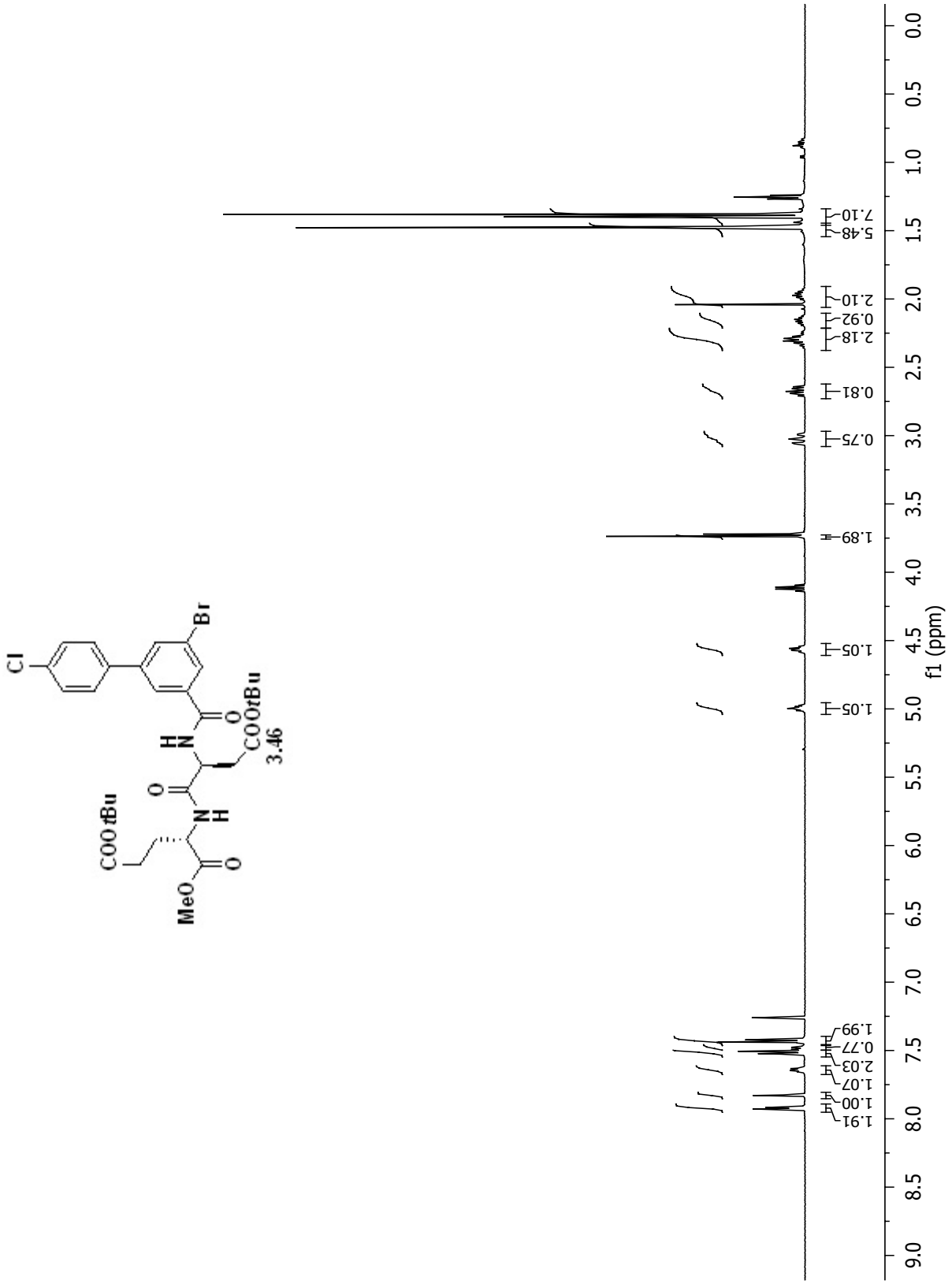

N
N
HO
3.40g
0
10
20
30
40
50
60
70
80
90
100
110
120
130
140
150
160
170
180
190
200
f1 (ppm)

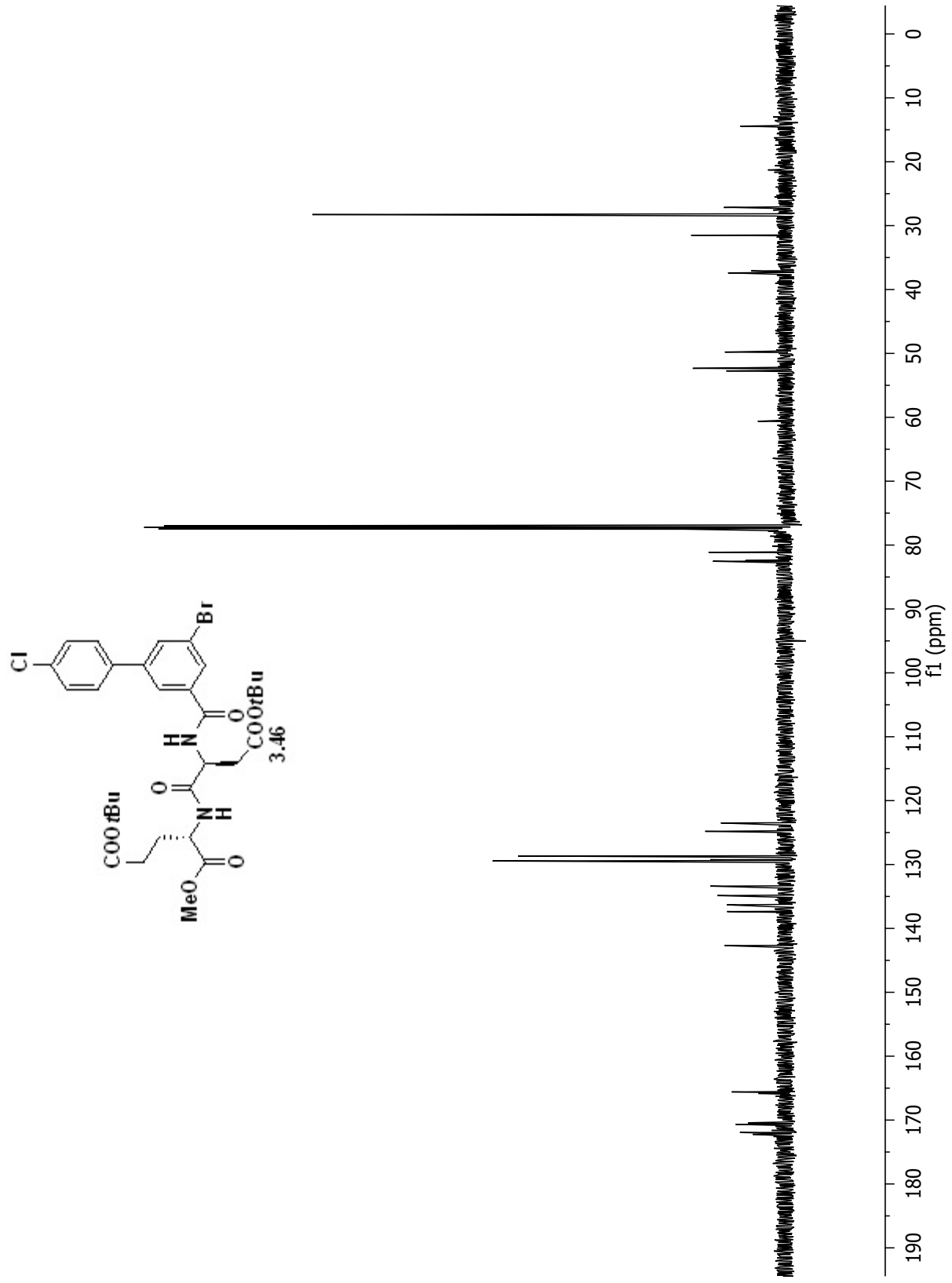

Cl
O
N
H
N
O
O
COOtBu
COOtBu
MeO
O
O
3.42
10.0
9.5
9.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
f1 (ppm)
1.00
0.98
1.03
0.93
0.90
4.84
2.12
1.12
1.08
0.96
1.06
1.07
2.03
1.43
2.76
2.24
1.07
1.01
1.83
1.01
1.03
10.15
9.40

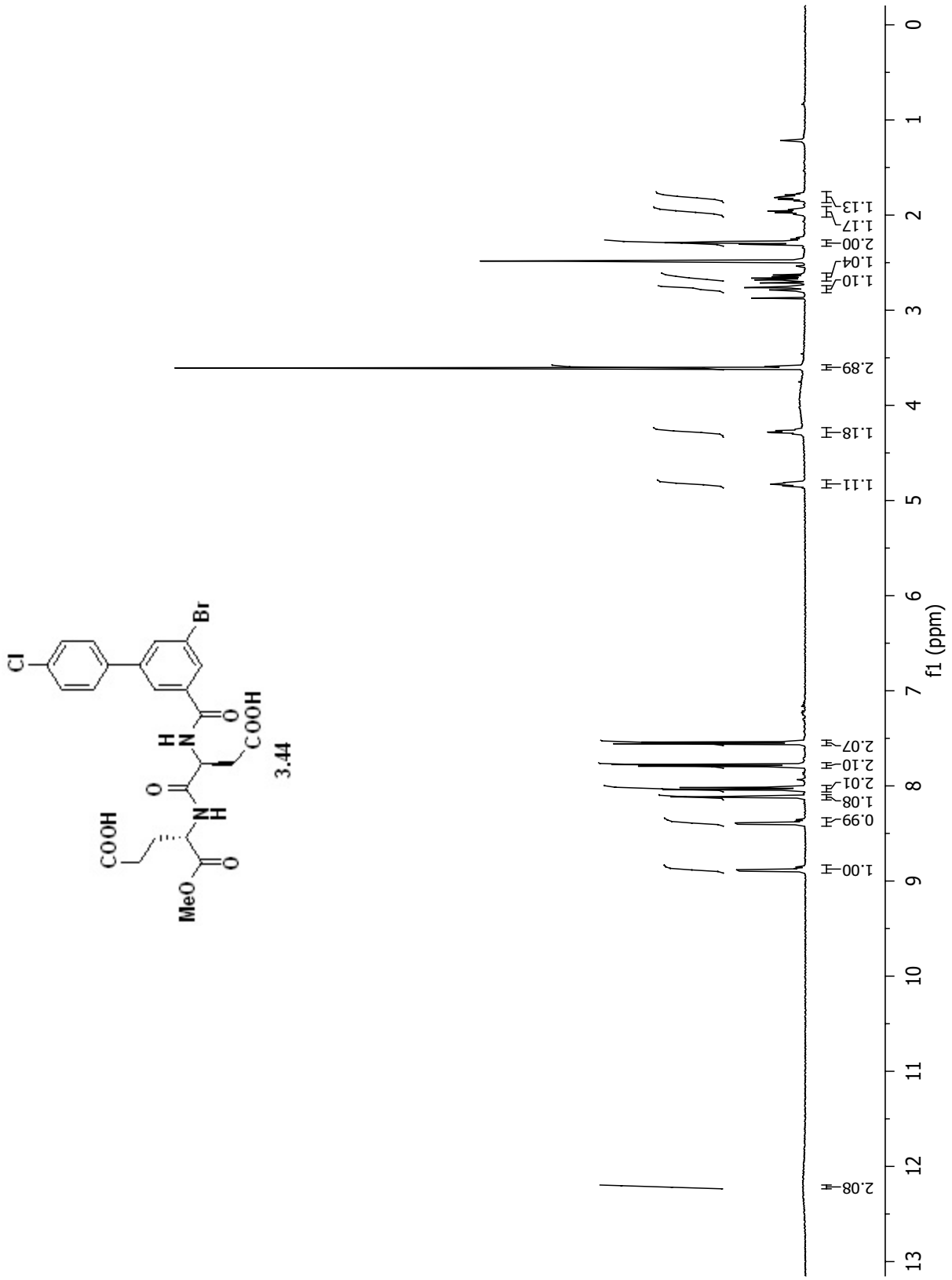

Cl
COOtBu
MeO
COOtBu
COOtBu
3.42
f1 (ppm)

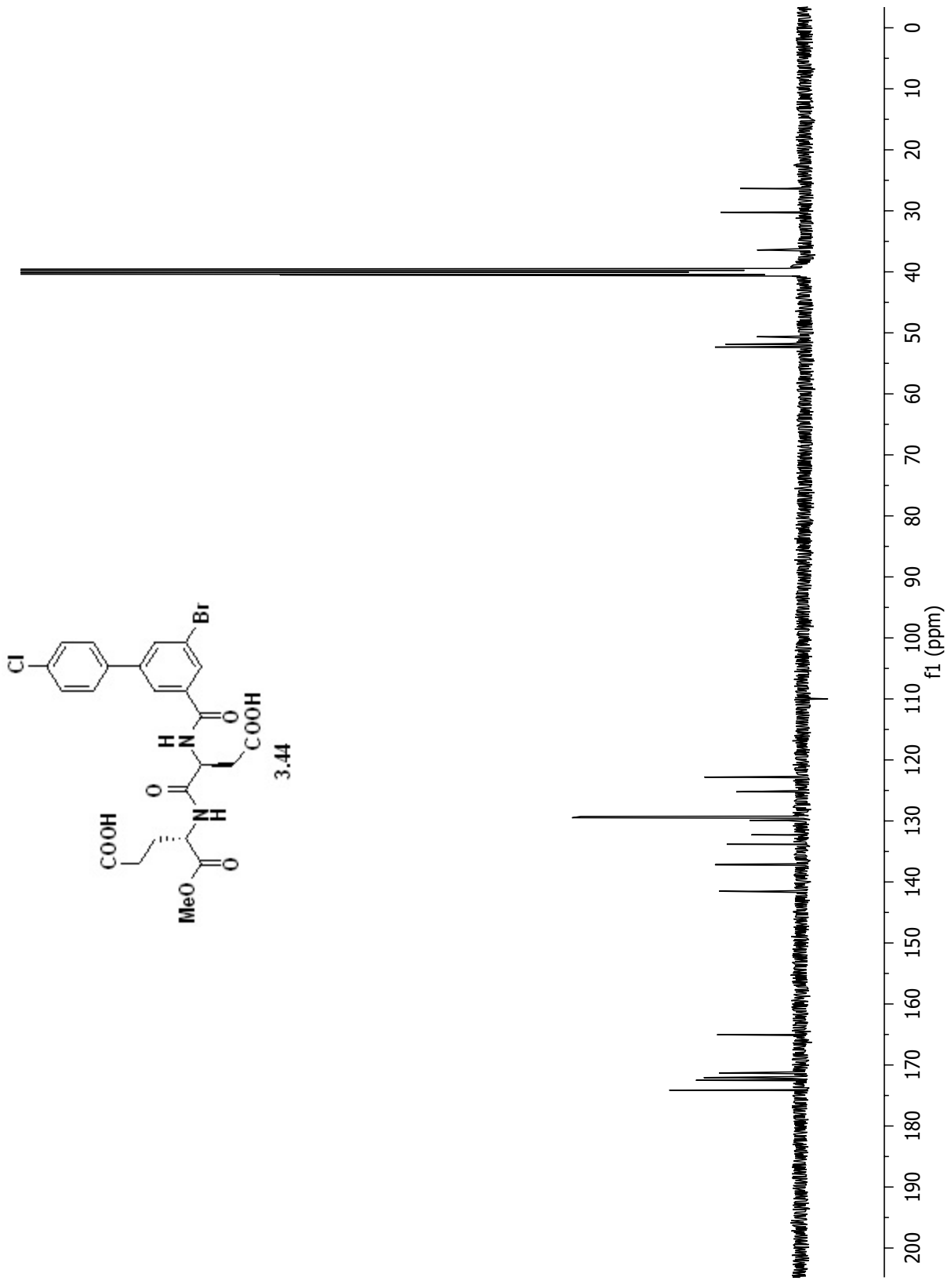

Cl
COOH
MeO
COOH
3.43
CF3COO
HN
1.00
1.01
0.83
0.90
1.06
3.27
3.19
1.04
2.00
3.95
0.83
1.59
0.94
2.62
2.10
0.94
0.63
3.66
1.03
0.60
f1 (ppm)

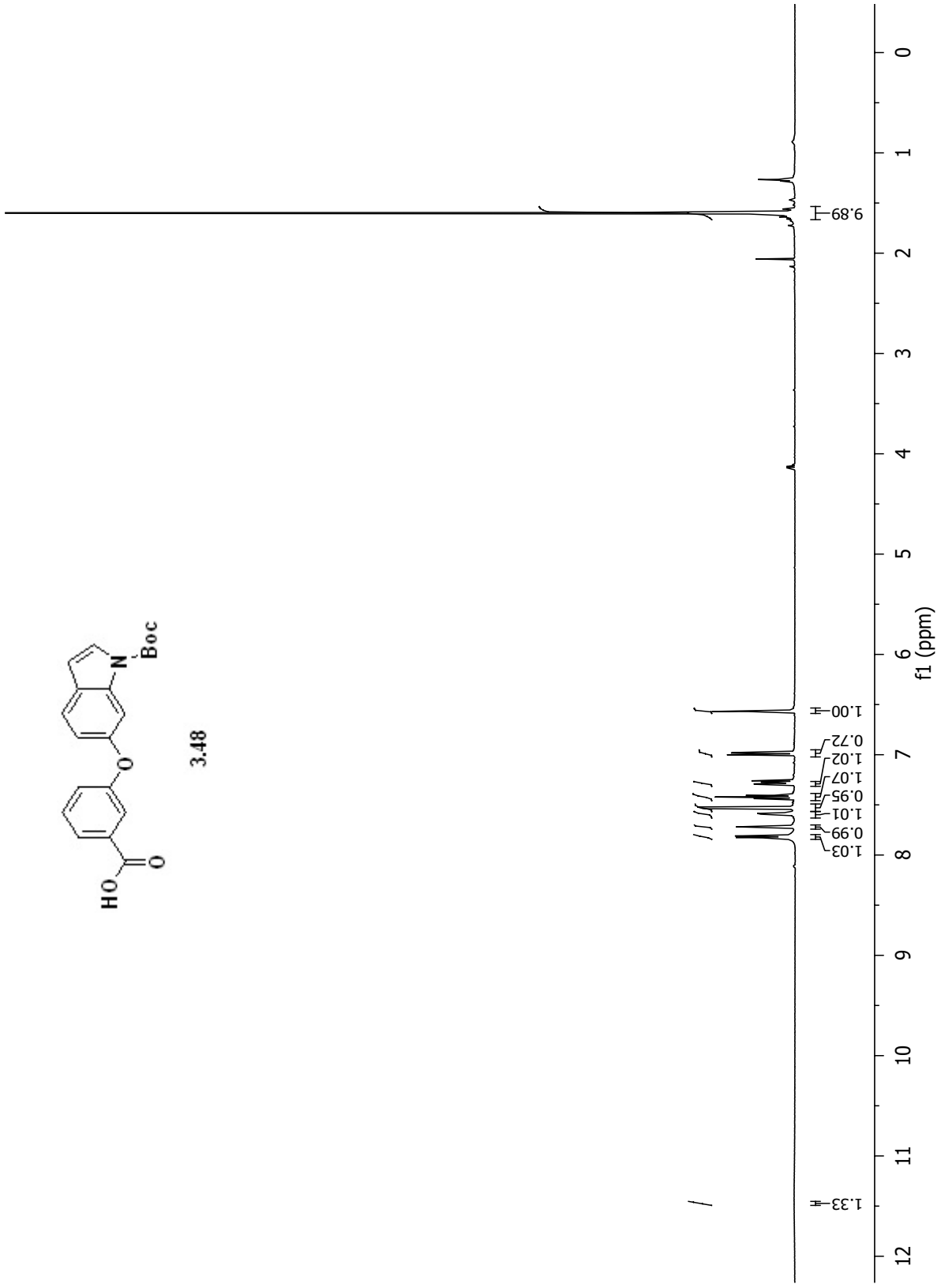

Cl
O
N
H
O
COOH
3.43
COOH
N
H
O
MeO
O
N
N
HN
$CF_3COO^{\ominus}$
f1 (ppm)
200
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0

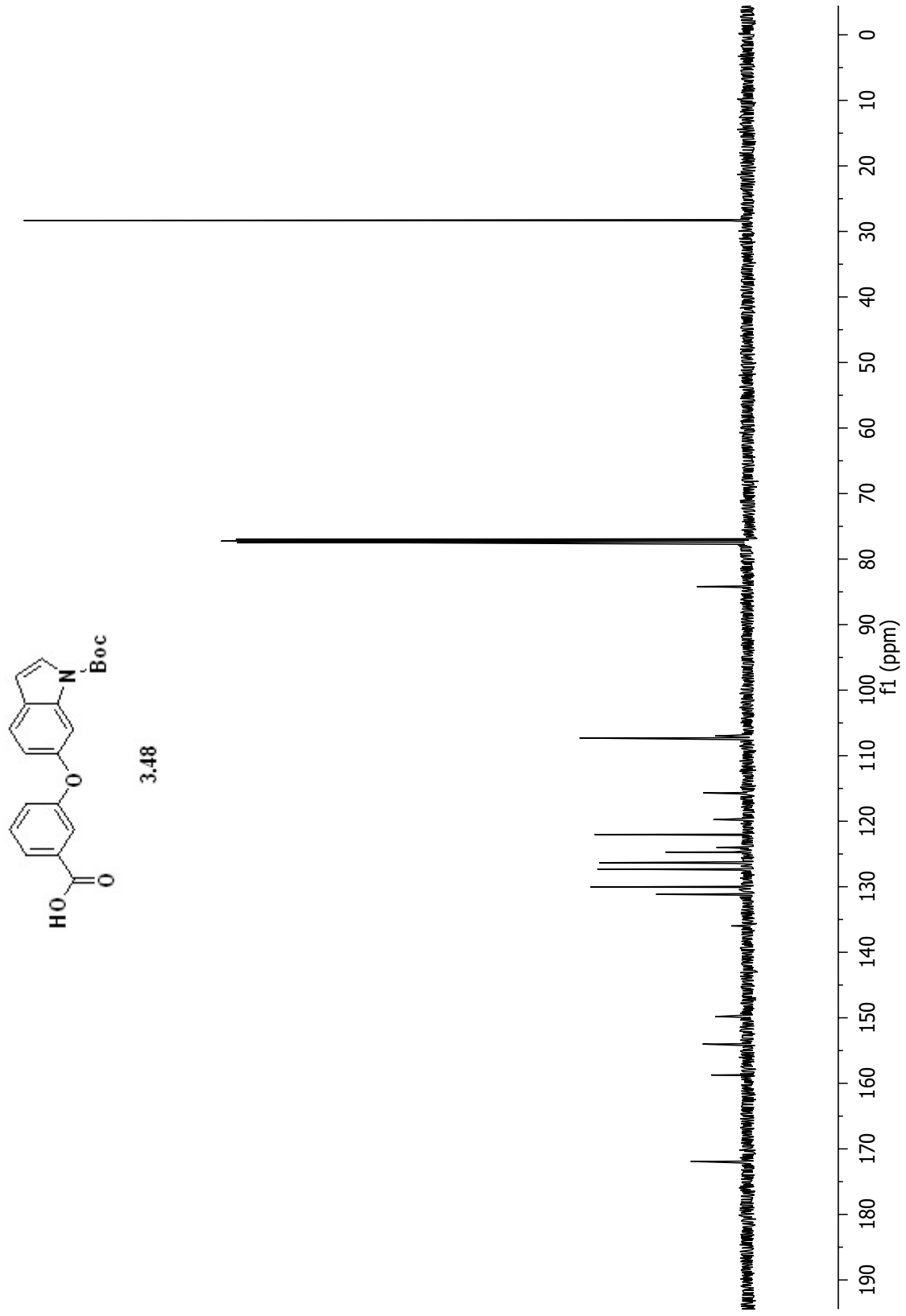

Cl
Br
COOtBu
O
H
N
O
COOtBu
3.46
O
H
N
COOtBu
O
MeO
O
9.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
f1 (ppm)
1.91
1.00
1.07
2.03
0.77
1.99
1.05
1.05
1.89
0.75
0.81
2.18
0.92
2.10
5.48
7.10

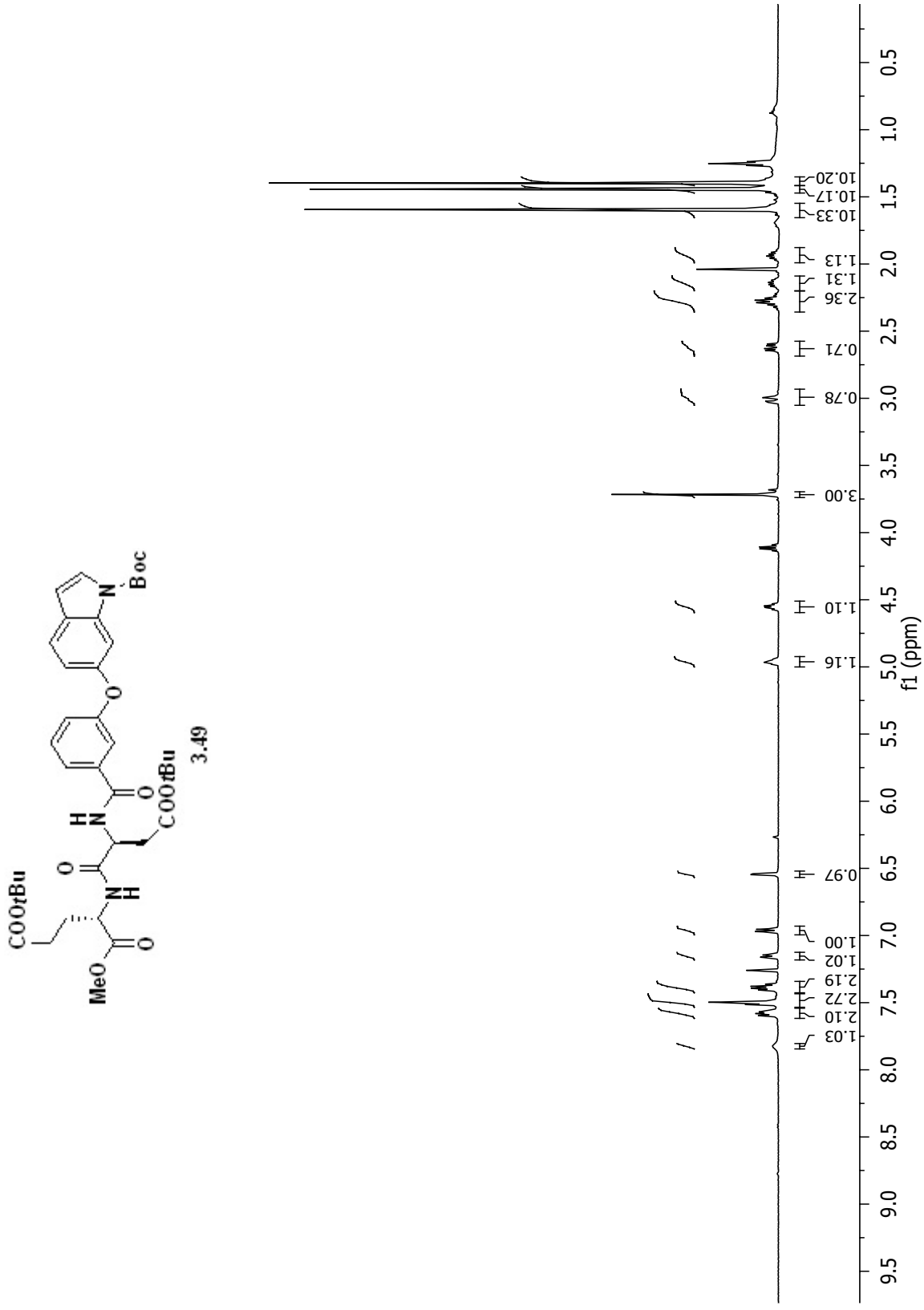

Cl
Br
O
H
N
O
COOtBu
3.46
O
H
N
COOtBu
O
MeO
O
0
10
20
30
40
50
60
70
80
90
100
110
120
130
140
150
160
170
180
190
f1 (ppm)

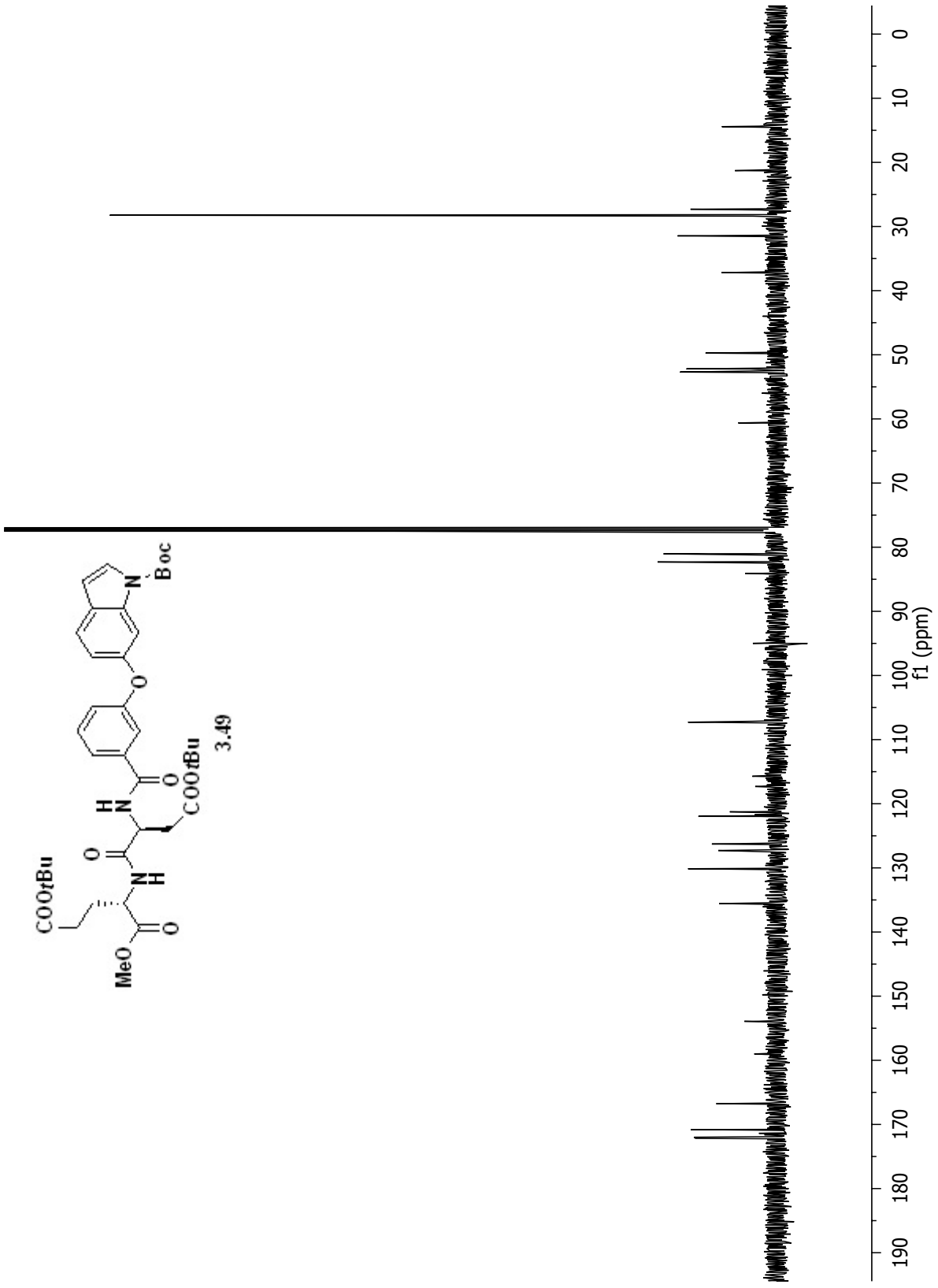

Cl
Br
O
H
N
COOH
O
O
H
N
COOH
O
MeO
3.44
2.08
1.00
0.99
1.08
2.01
2.10
2.07
1.11
1.18
2.89
1.10
1.04
2.00
1.17
1.13
0
1
2
3
4
5
6
7
8
9
10
11
12
13
f1 (ppm)

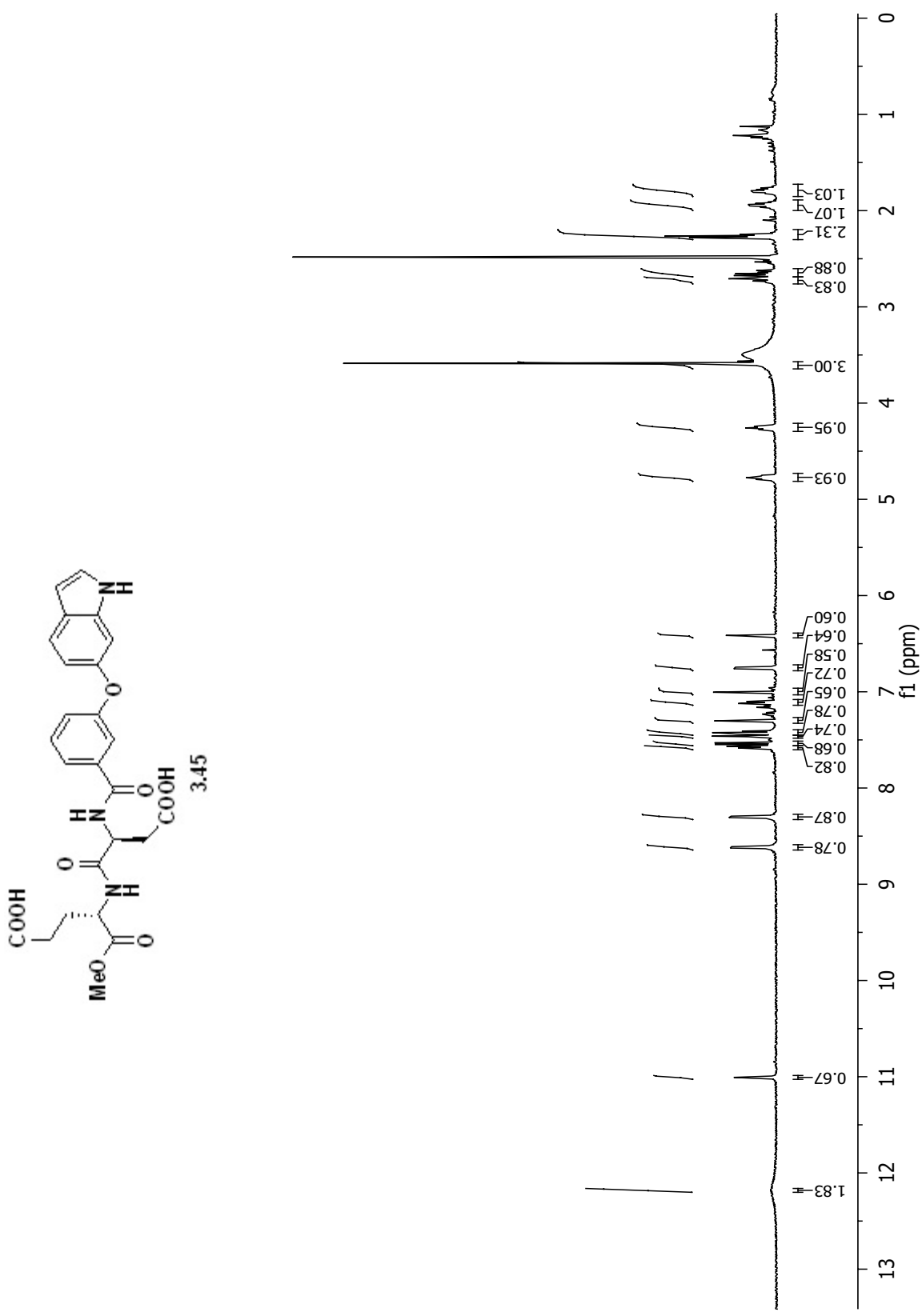

Cl
Br
O
H
N
COOH
O
O
H
N
COOH
O
MeO
3.44
f1 (ppm)
0
10
20
30
40
50
60
70
80
90
100
110
120
130
140
150
160
170
180
190
200

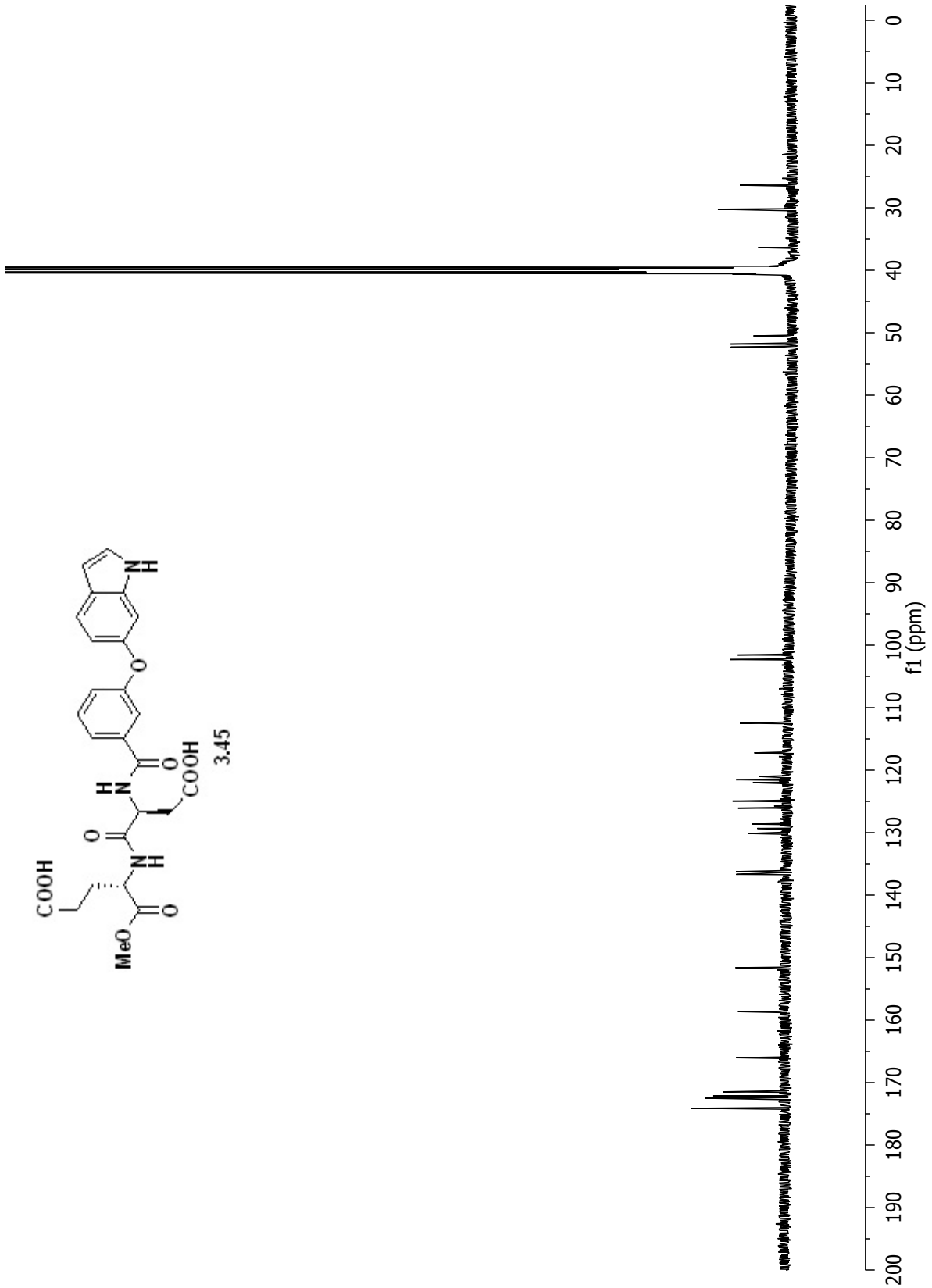

HO
O
O
O
N
Boc
3.48
9.89
1.00
1.03
0.99
1.01
0.95
1.07
1.02
0.72
1.33
f1 (ppm)

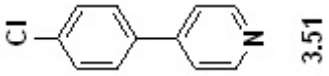

HO
O
O
O
N
Boc
3.48
f1 (ppm)


COOtBu
MeO
O
O
H
N
O
H
N
O
COOtBu
O
O
N
Boc
3.49
f1 (ppm)
9.5
9.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
1.03
2.10
2.72
2.19
1.02
1.00
0.97
1.16
1.10
3.00
0.78
0.71
2.36
1.31
1.13
10.33
10.17
10.20

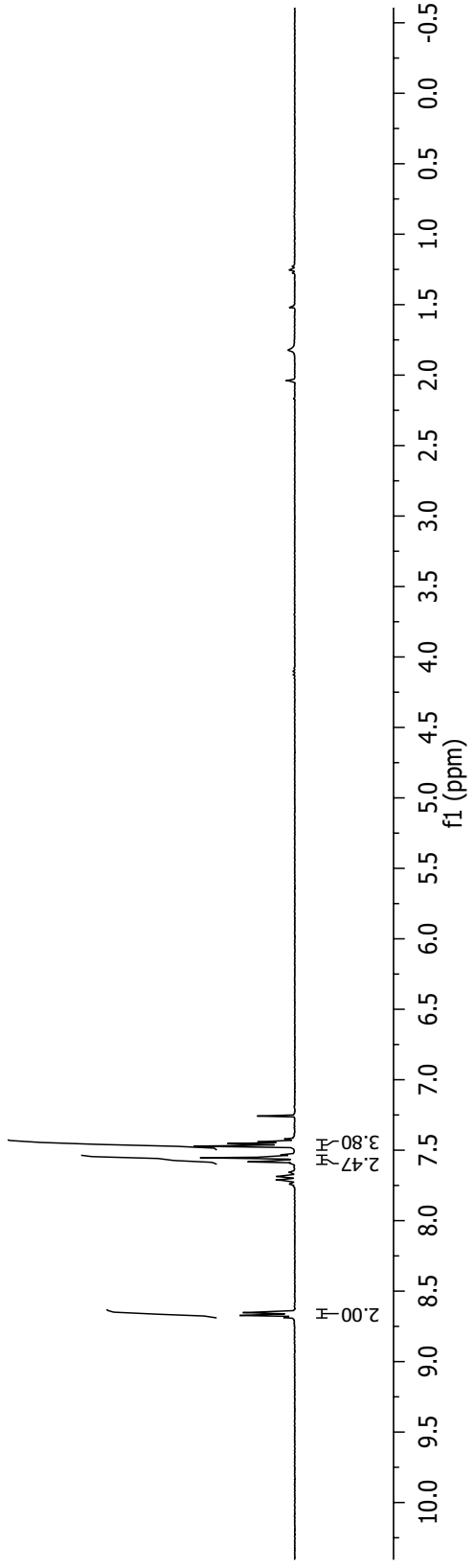

COOtBu
MeO
O
N
H
O
O
N
H
O
COOtBu
O
O
N
Boc
3.49
f1 (ppm)
0
10
20
30
40
50
60
70
80
90
100
110
120
130
140
150
160
170
180
190

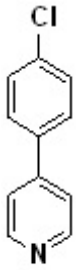

COOH
MeO
O
O
H
N
O
H
N
O
O
O
COOH
O
H
N
3.45
1.83
0.67
0.78
0.87
0.82
0.68
0.74
0.78
0.65
0.72
0.58
0.64
0.60
0.93
0.95
3.00
0.83
0.88
2.31
1.07
1.03
0
1
2
3
4
5
6
7
8
9
10
11
12
13
f1 (ppm)


COOH
MeO
O
O
H
N
O
H
N
O
O
COOH
O
H
N
3.45
f1 (ppm)
0
10
20
30
40
50
60
70
80
90
100
110
120
130
140
150
160
170
180
190
200

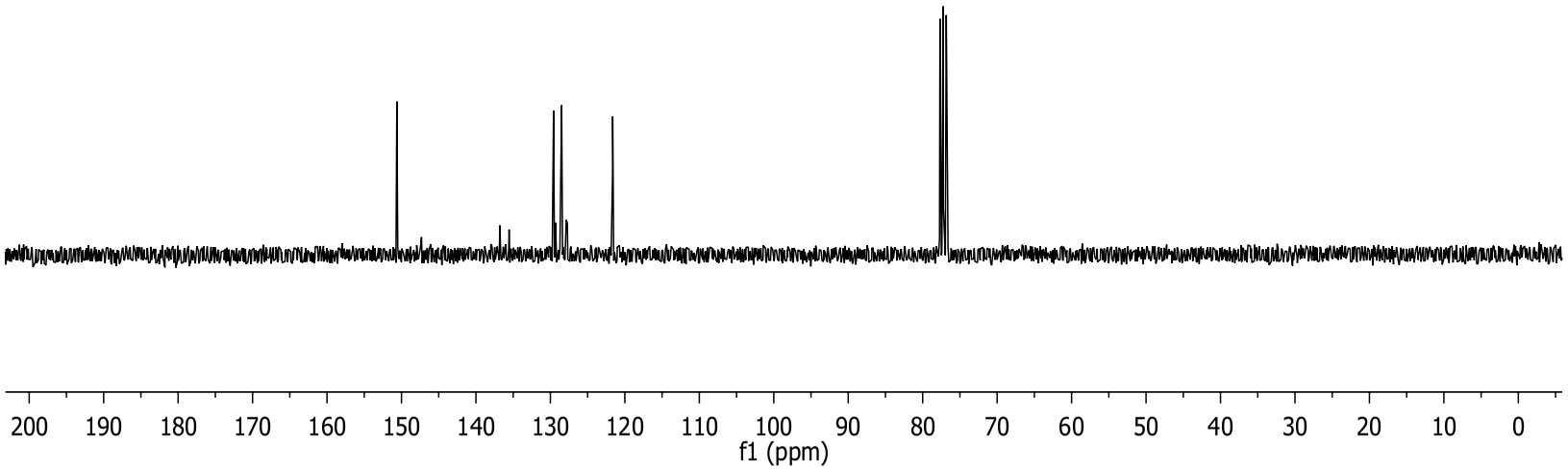

Cl
N

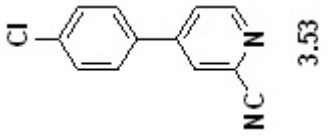
3.51


2.00
3.80
2.47
10.0
9.5
9.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
-0.5
f1 (ppm)

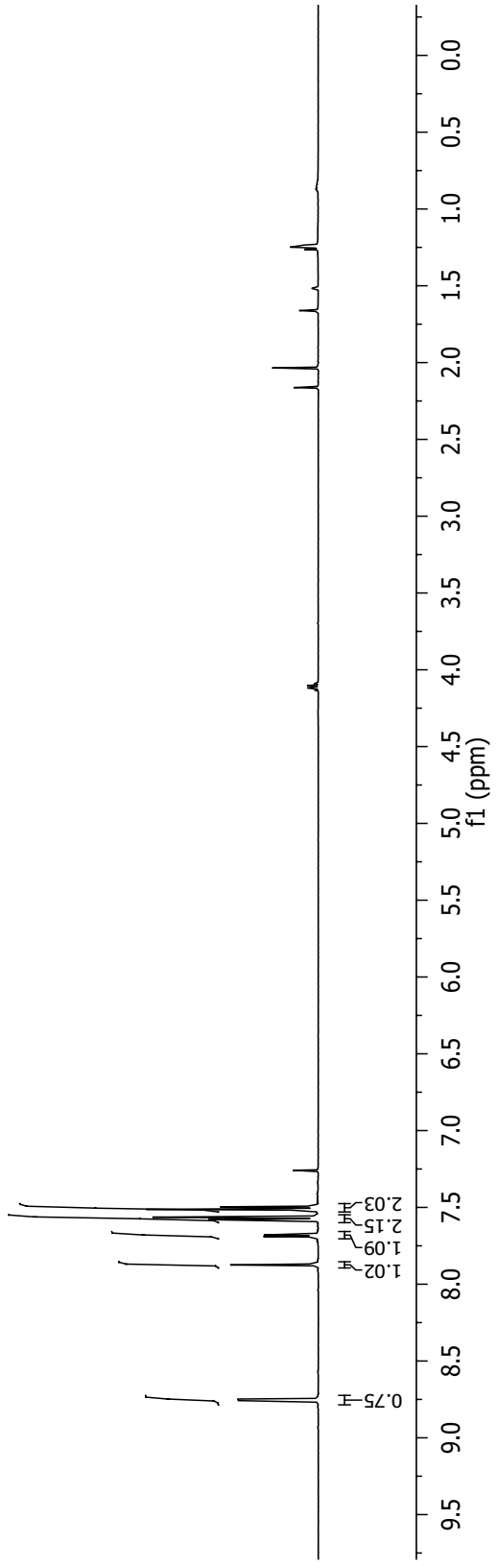

Cl
N

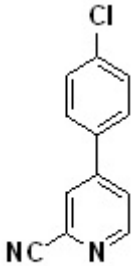
3.51


200
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
f1 (ppm)

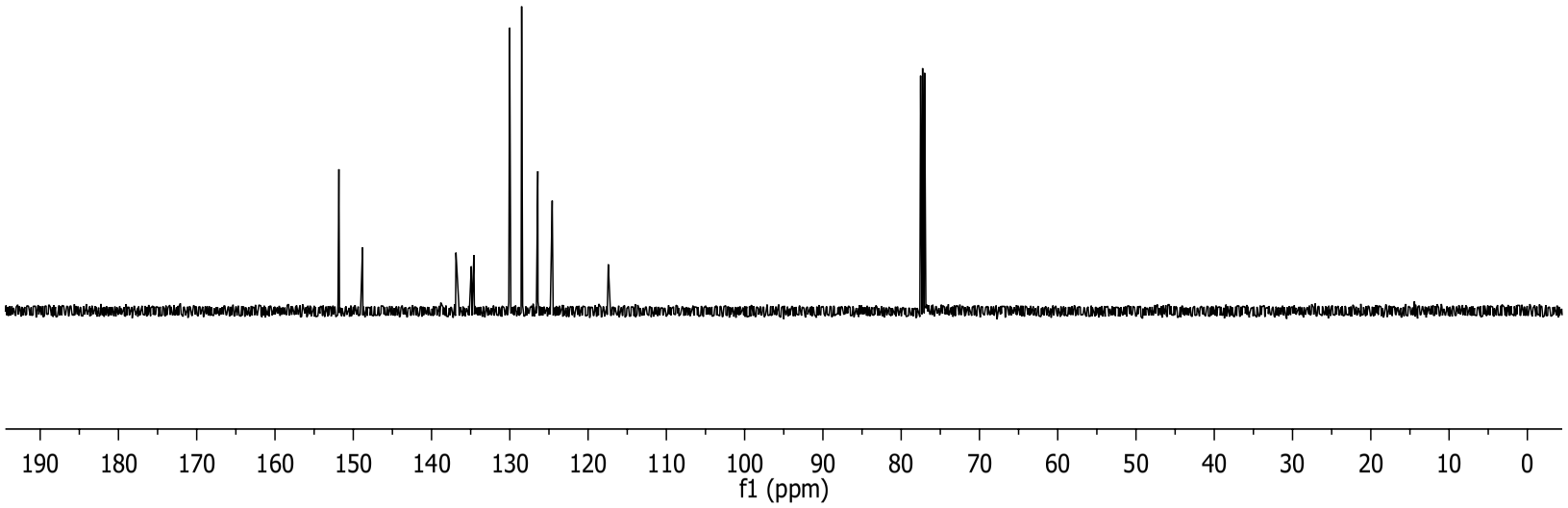

Cl
N
NC

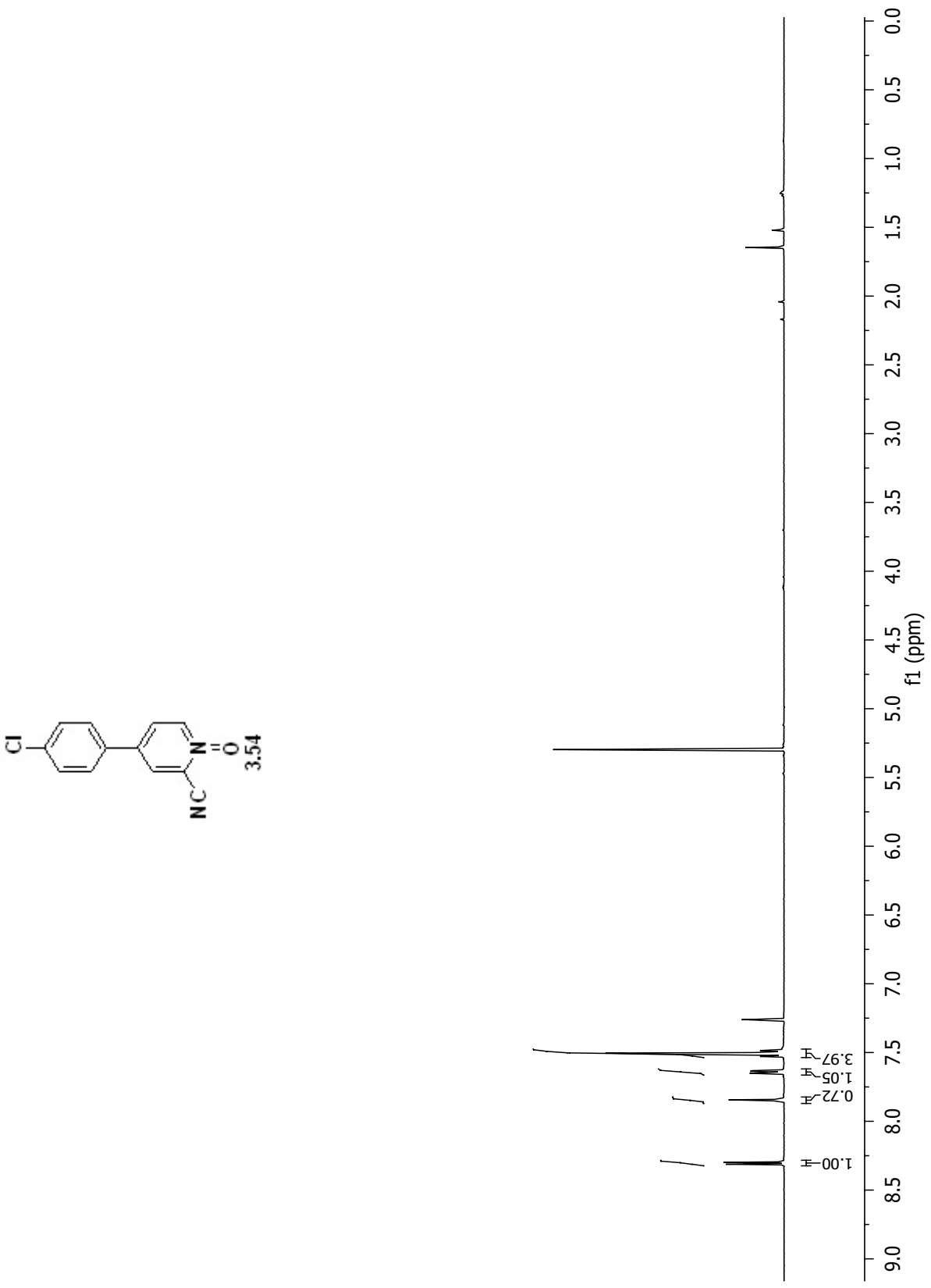
3.53

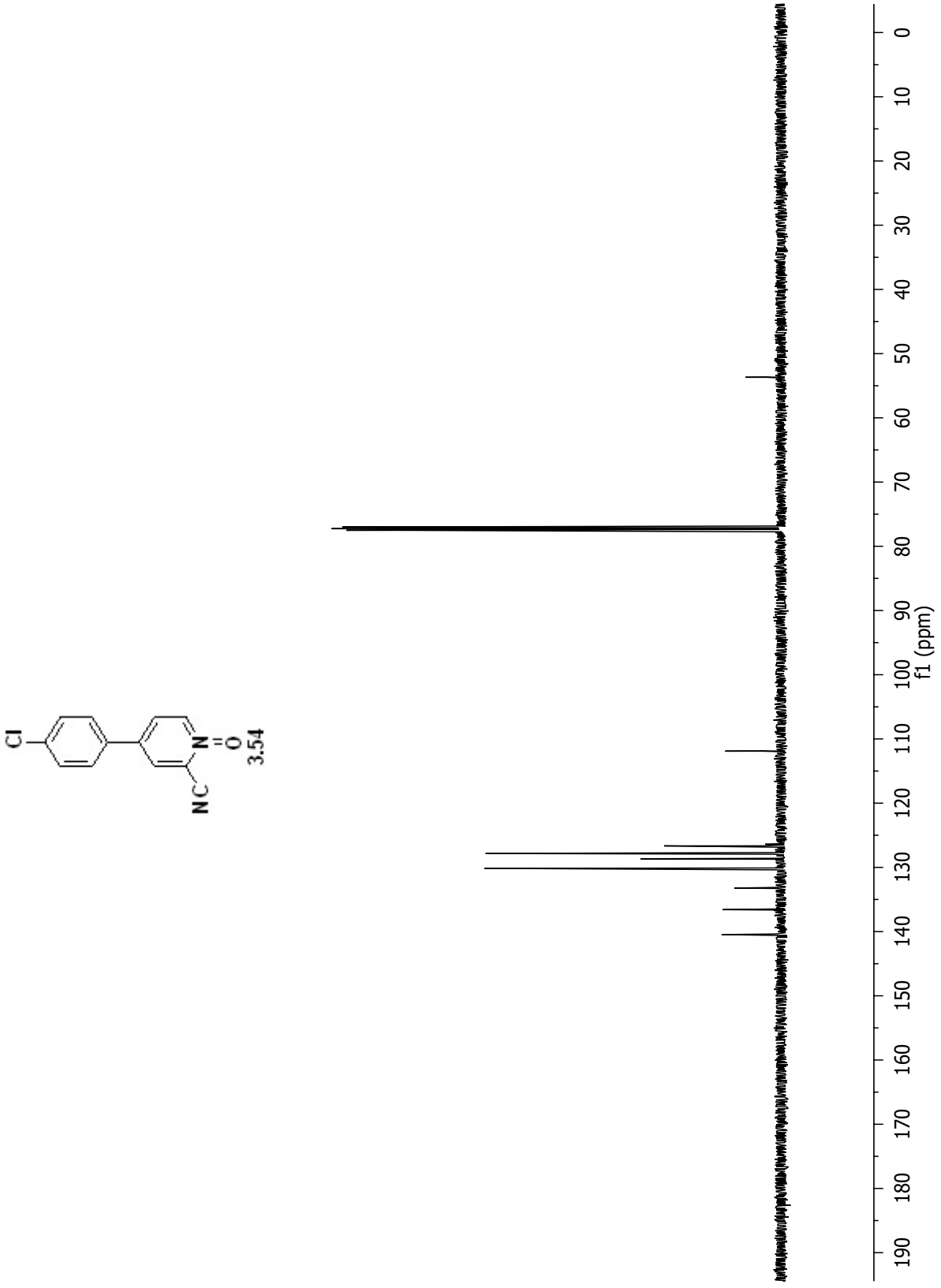

0.75
1.02
1.09
2.15
2.03
9.5
9.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
f1 (ppm)

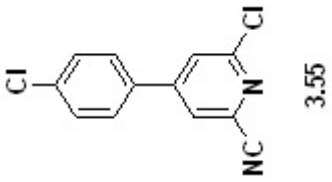

Cl
NC
N

3.53

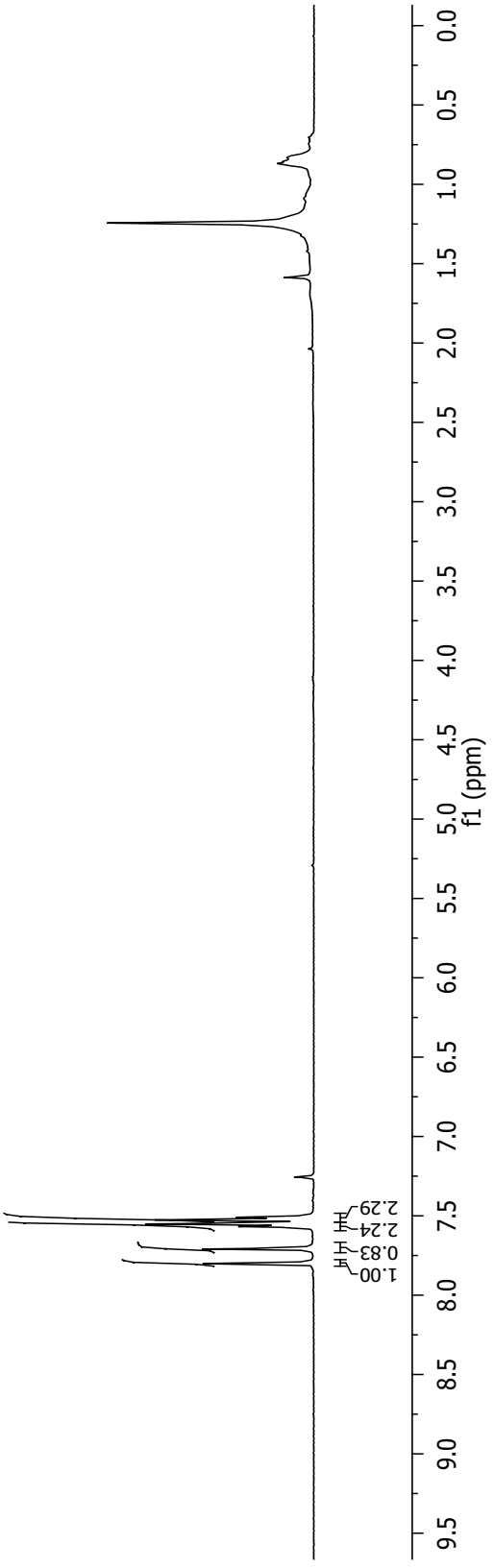

190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
f1 (ppm)

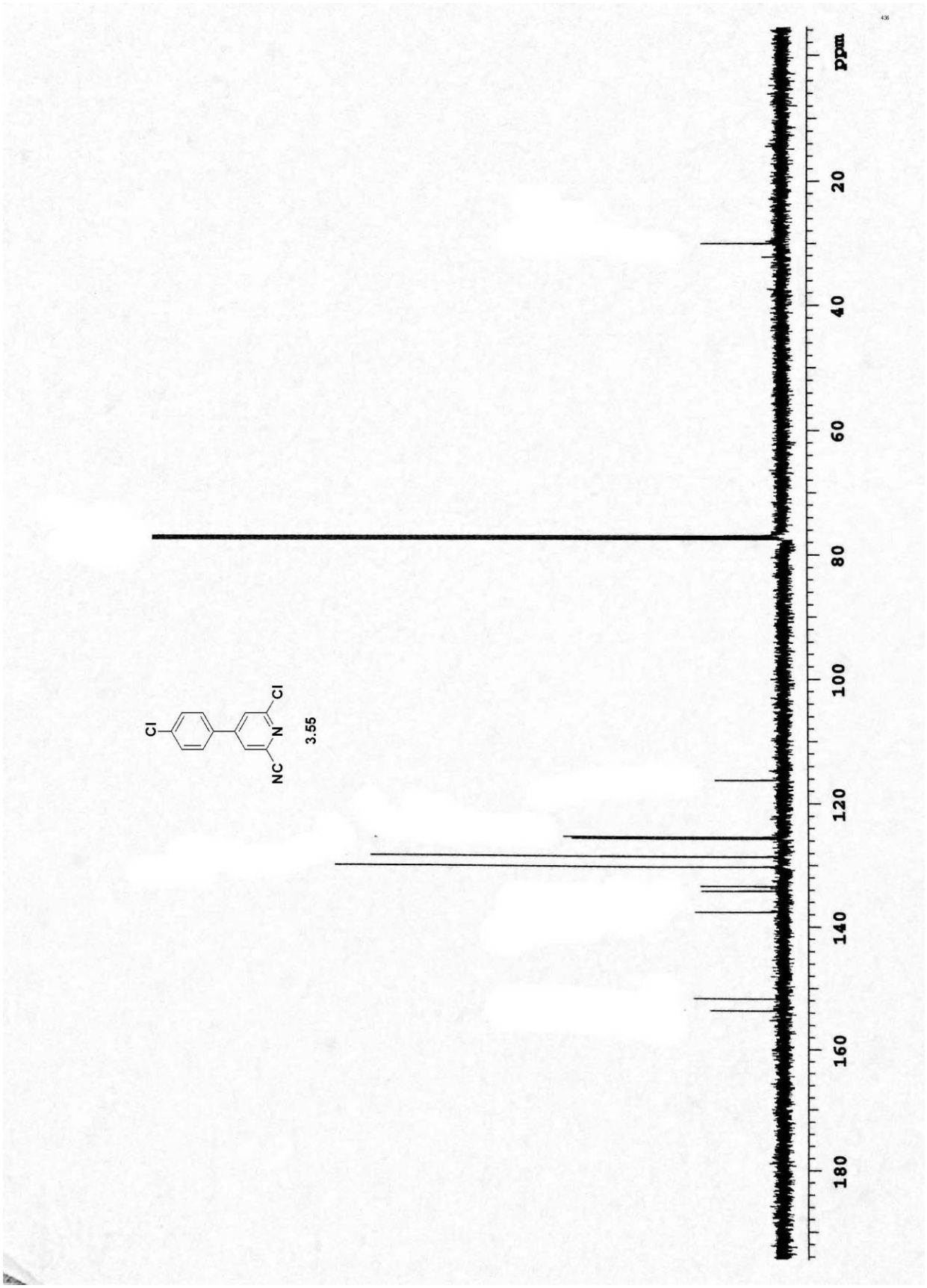

Cl
N
O
NC
3.54
1.00
0.72
1.05
3.97
9.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
f1 (ppm)

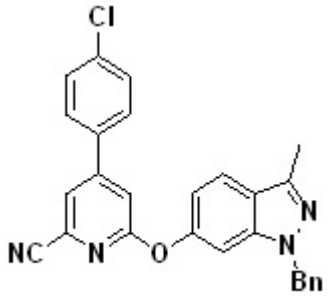

Cl
N
O
NC
3.54
0
10
20
30
40
50
60
70
80
90
100
110
120
130
140
150
160
170
180
190
f1 (ppm)


Cl
Cl
N
NC

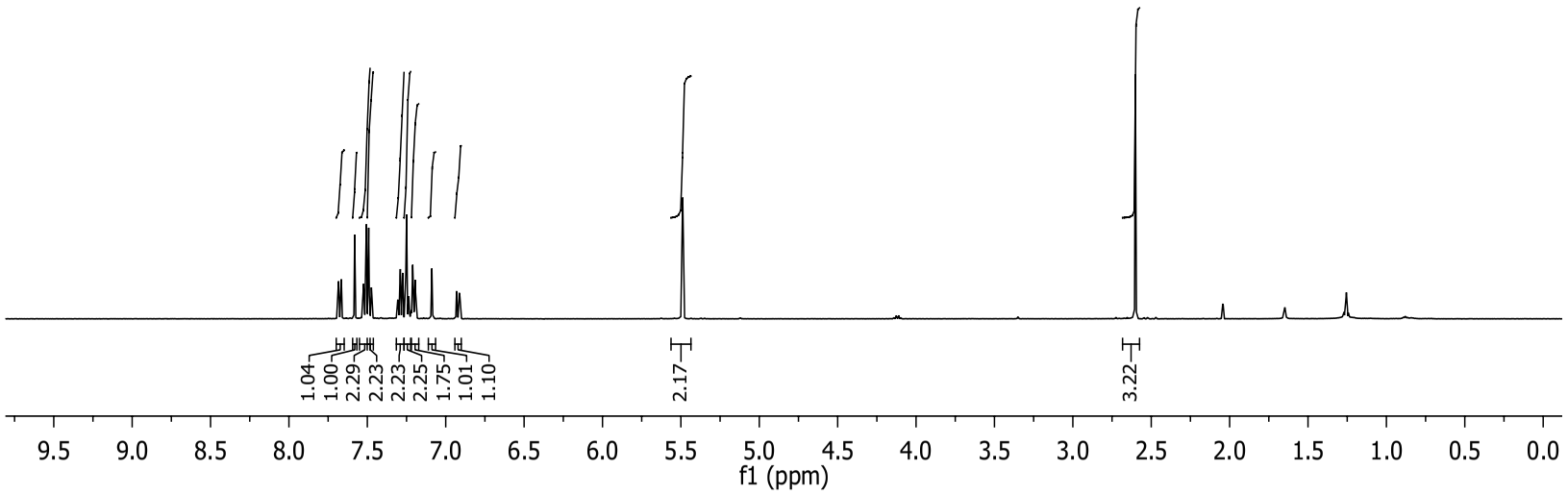
3.55

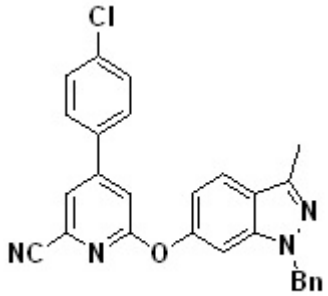

1.00
0.83
2.24
2.29
9.5
9.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
f1 (ppm)


Cl
Cl
N
NC
3.55
ppm
180
160
140
120
100
80
60
40
20

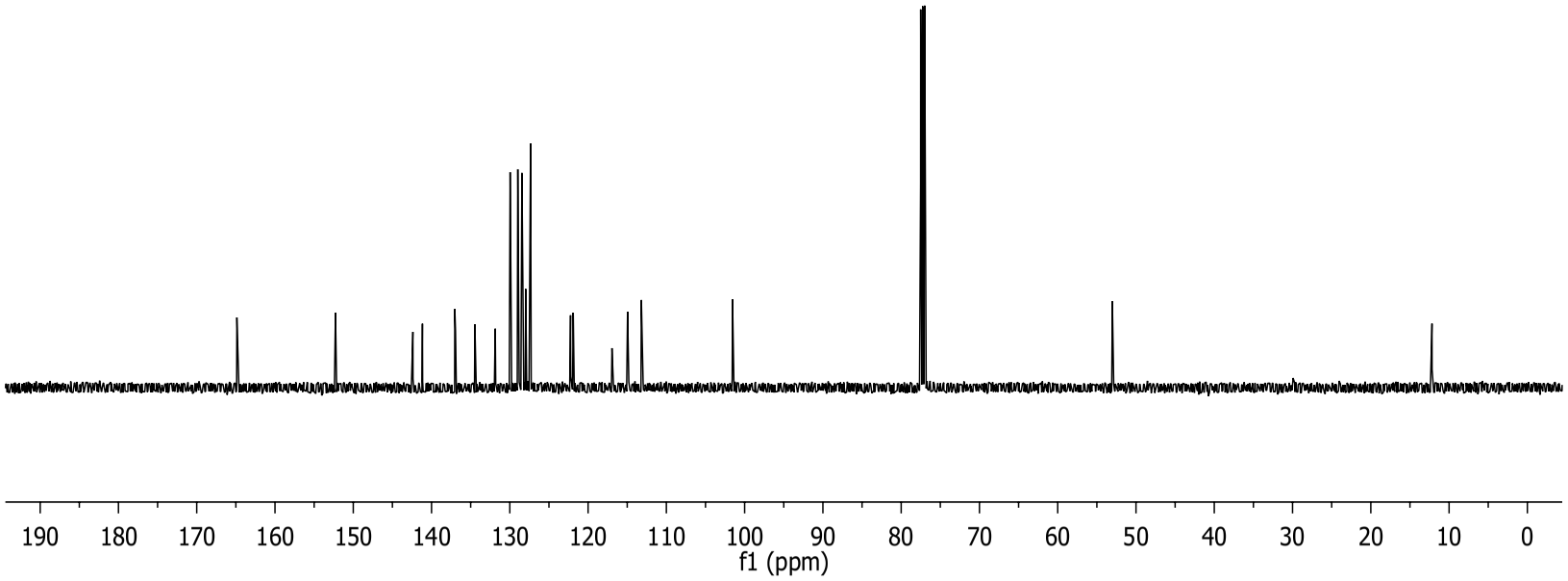

Cl
NC
N
O
N
N
Bn

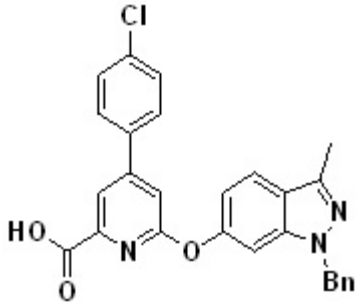
3.56


1.04
1.00
2.29
2.23
2.23
2.25
1.75
1.01
1.10
2.17
3.22
9.5
9.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
f1 (ppm)

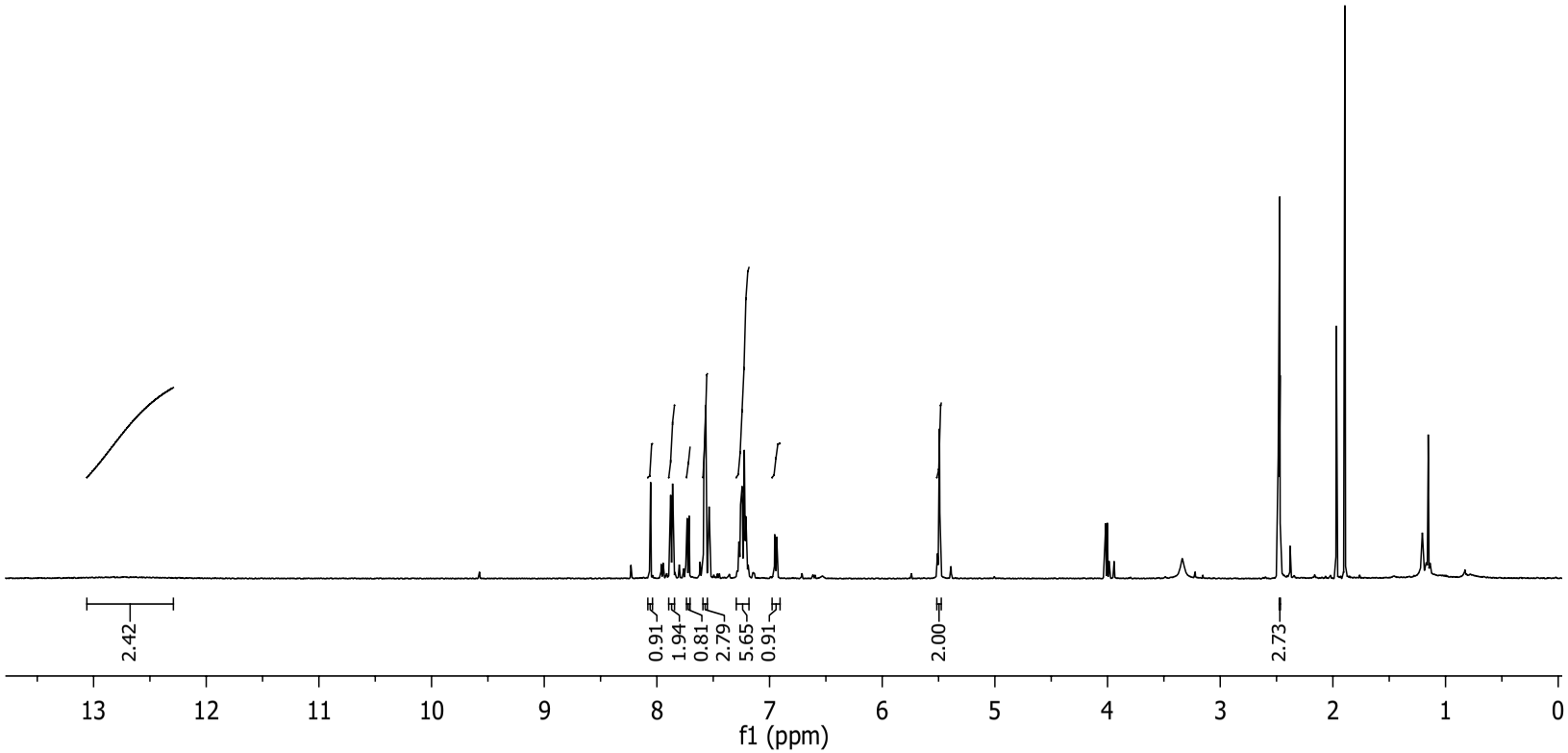

Cl
NC
N
O
N
N
Bn

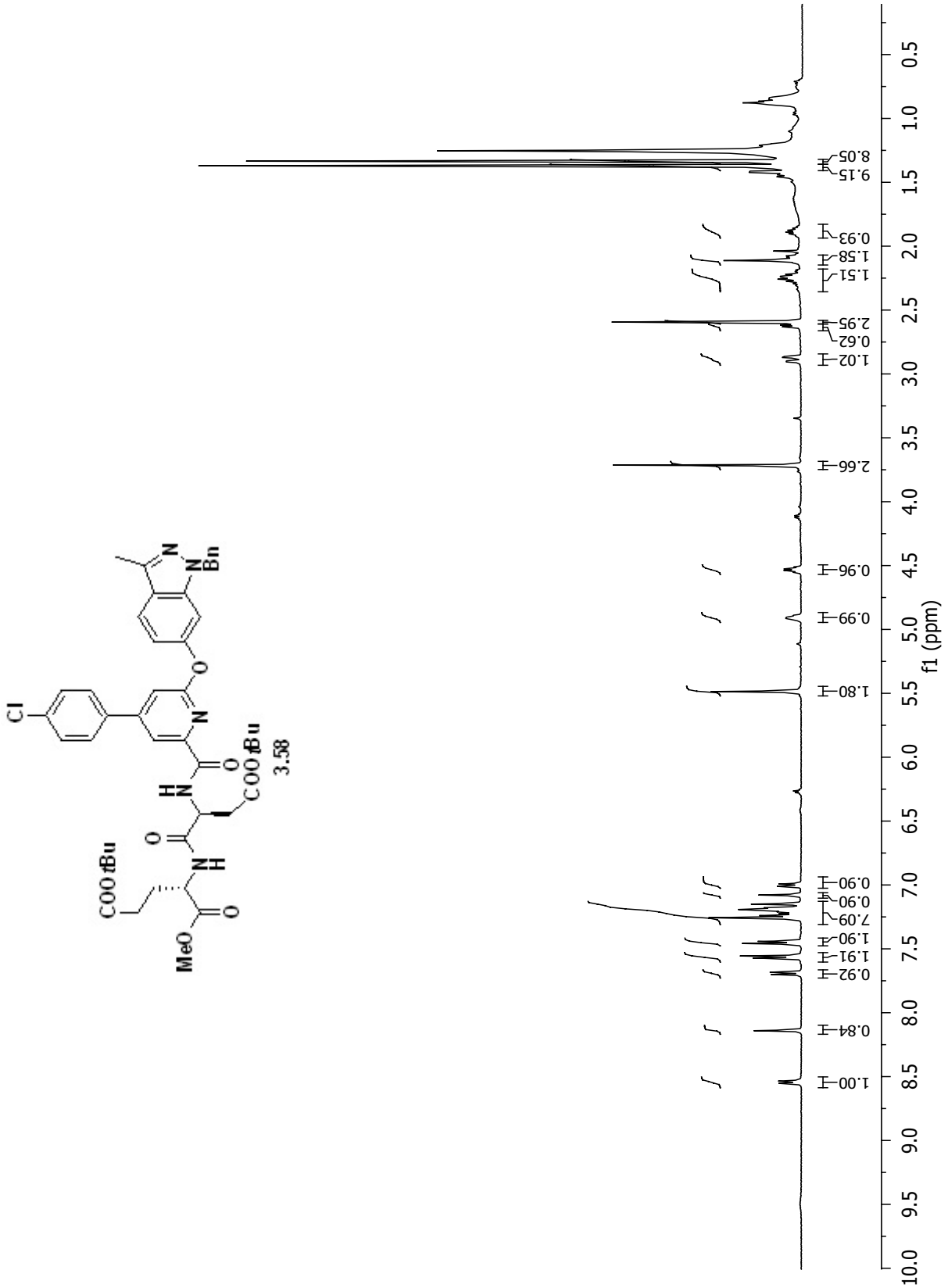
3.56

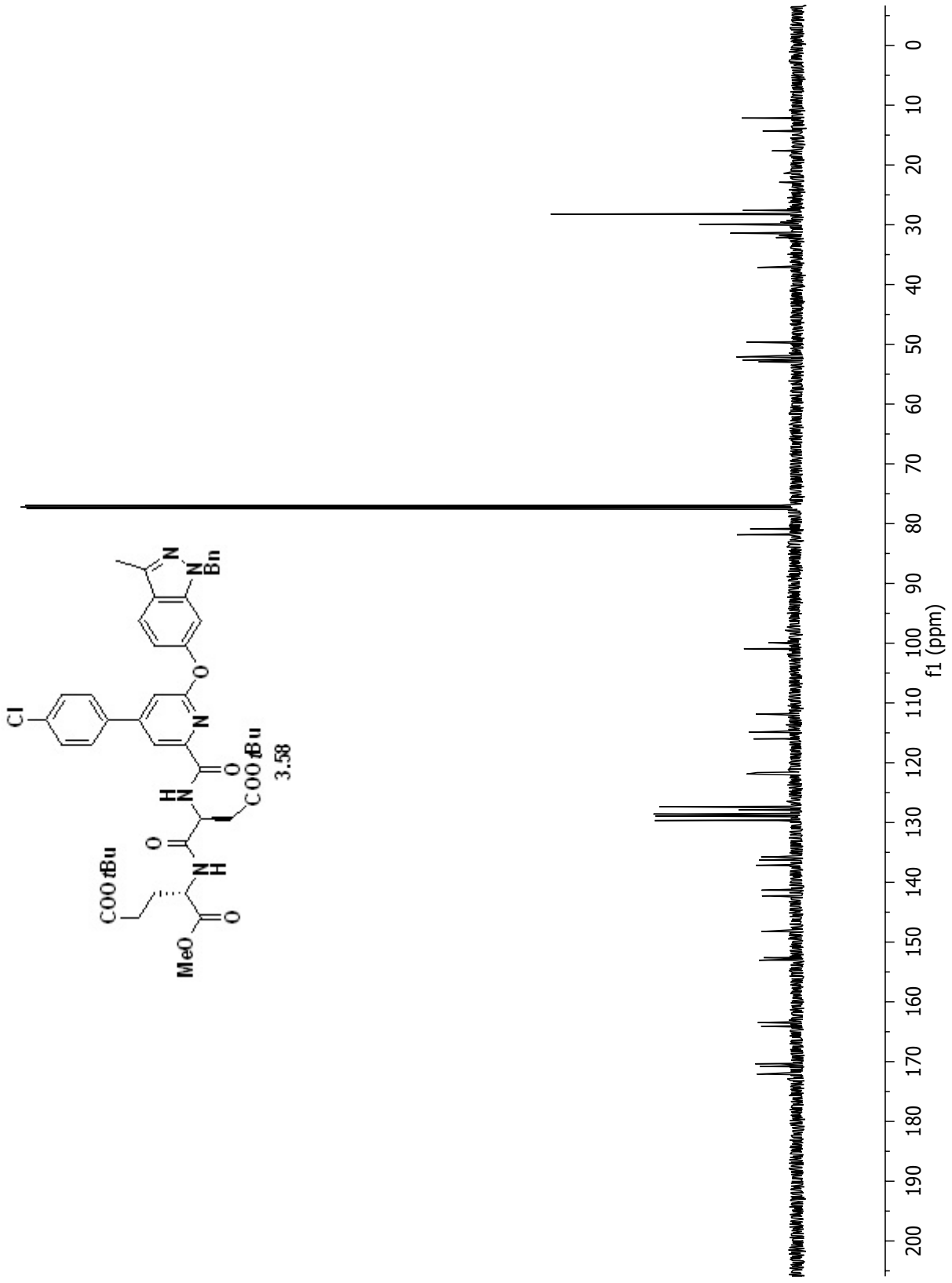

190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
f1 (ppm)

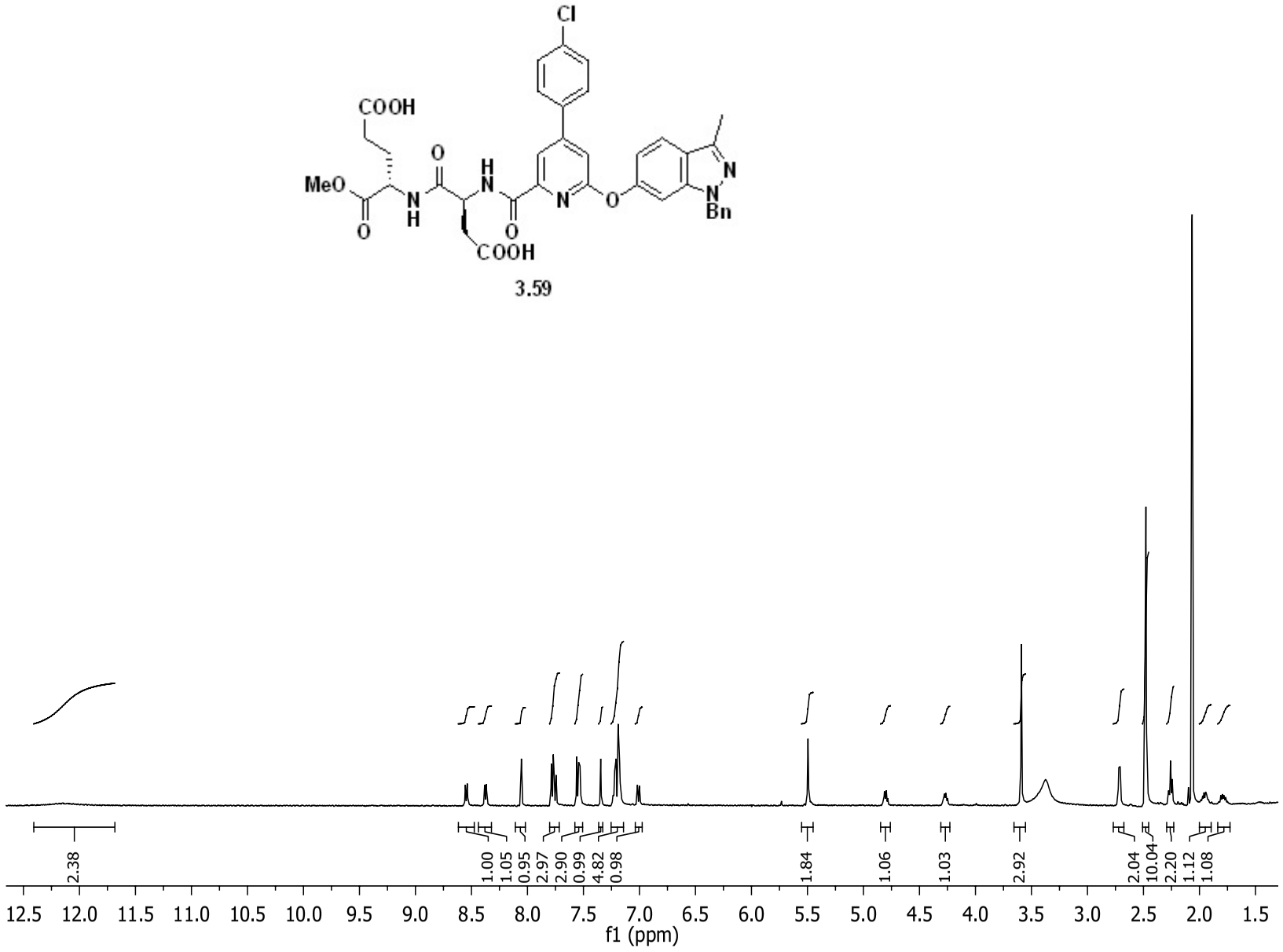

Cl
HO
O
N
O
N
N
Bn

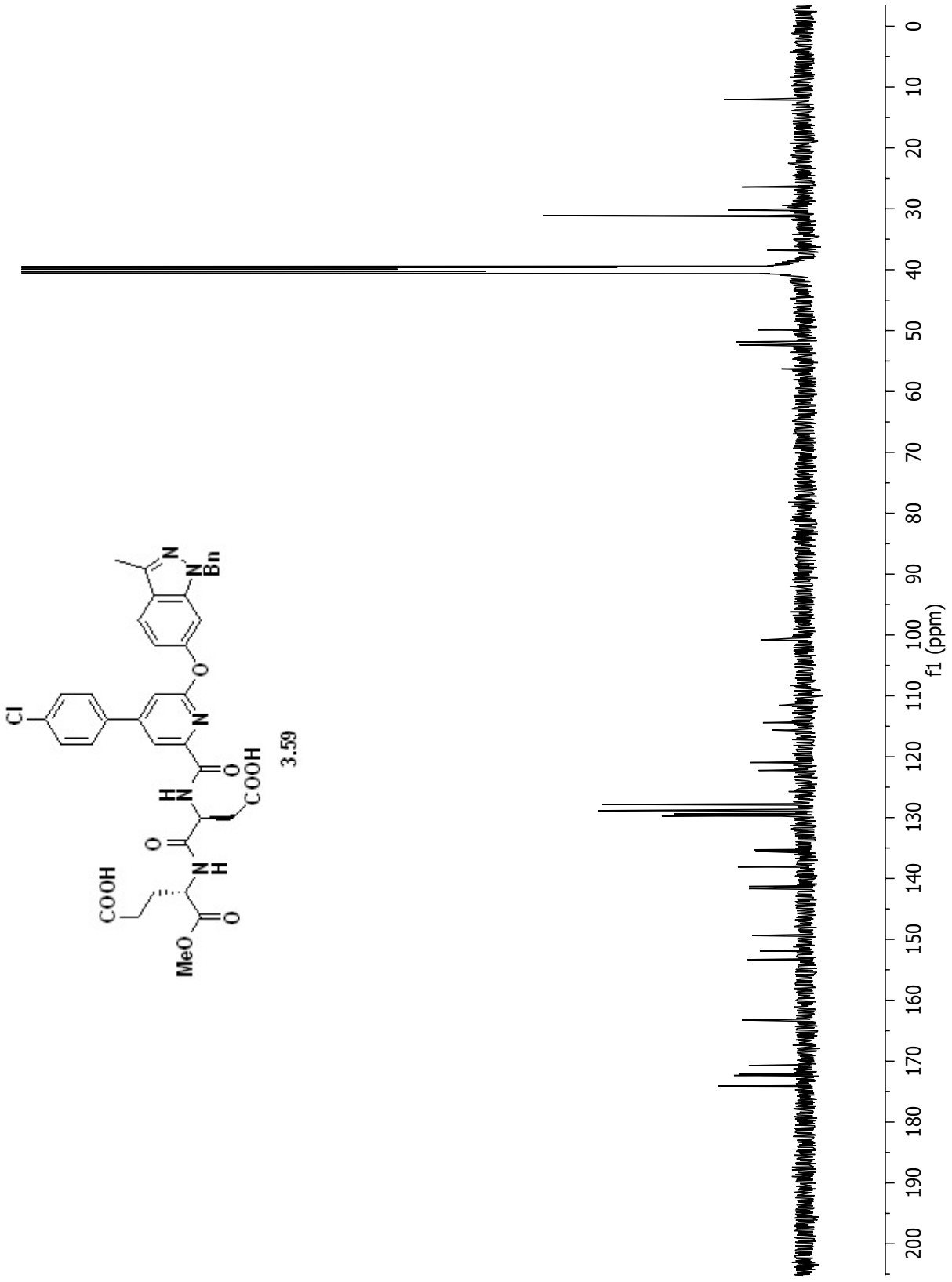
3.57


2.42
0.91
1.94
0.81
2.79
5.65
0.91
2.00
2.73
13
12
11
10
9
8
7
6
5
4
3
2
1
0
f1 (ppm)


Cl
N
O
N
Bn
O
H
N
O
COOtBu
3.58
H
N
O
COOtBu
O
MeO
10.0
9.5
9.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
f1 (ppm)
1.00
0.84
0.92
1.91
1.90
7.09
0.90
0.90
1.80
0.99
0.96
2.66
1.02
0.62
2.95
1.51
1.58
0.93
9.15
8.05


Cl
N
O
N
Bn
O
H
N
O
COOtBu
3.58
O
H
N
COOtBu
O
MeO
O
200
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
f1 (ppm)


Cl
COOH
MeO
O
H
N
N
O
N
N
Bn
O
COOH
3.59
2.38
1.00
1.05
0.95
2.97
2.90
0.99
4.82
0.98
1.84
1.06
1.03
2.92
2.04
10.04
2.20
1.12
1.08
12.5 12.0 11.5 11.0 10.5 10.0 9.5 9.0 8.5 8.0 7.5 7.0 6.5 6.0 5.5 5.0 4.5 4.0 3.5 3.0 2.5 2.0 1.5
f1 (ppm)


Cl
N
O
N
Bn
O
H
N
O
COOH
3.59
O
H
N
COOH
O
MeO
f1 (ppm)
0
10
20
30
40
50
60
70
80
90
100
110
120
130
140
150
160
170
180
190
200